



Deutsches Zentrum
für Luft- und Raumfahrt

Otto-von-Guericke-Universität Magdeburg
Fakultät für Naturwissenschaften
Institut für Psychologie
Abteilung Umweltpsychologie

Erstgutachterin: Fr. Dr. Annika Dreßler
Zweitgutachterin: Fr. Prof. Dr. Ellen Matthies

Masterarbeit

„Bahn frei?“:

**Effektive Förderung des Kontrollblicks an Bahnübergängen und
Analyse der Objektivität der gewählten Untersuchungsmethode**

Vorgelegt von:

Jennifer Nöske

Noeske.Jennifer@gmx.de

Matrikelnummer 192378

Braunschweig, den 12.02.2021

Zusammenfassung

Die Sicherheit an Bahnübergängen hängt von dem Verhalten der Straßenverkehrsteilnehmenden ab, da diese dem Schienenverkehr Vorrang gewähren müssen. Die Auswirkungen von Unfällen sind vor allem bei ungeschützten Straßennutzenden wie zu Fuß Gehenden oder Fahrradfahrenden gravierend: Unfälle verursachen neben den persönlichen Auswirkungen durch die Todesfälle und Verletzungen erhebliche Kosten durch Sachschäden, Umweltschäden und Folgeverspätungen. Ein Fokus des Forschungsprojektes SAFER-LC lag auf der Förderung des sicheren Kontrollblickverhaltens vor der Querung bei dieser Zielgruppe. In diesem Rahmen wurden zwei innovative, kostengünstige Maßnahmen an einem nicht-technisch gesicherten Bahnübergang getestet. Von der Entwicklungsphase bis hin zur post-hoc Evaluation der Maßnahmen wurde beachtet, dass durch die Umweltgestaltung sicherheitsrelevante Verhaltensweisen gefördert werden. Nach einer Kontrollerhebung wurden für jeweils zwei Wochen eine Straßenmarkierung mit dem Schriftzug *Kommt ein Zug?* aufgebracht sowie ein Blinklicht mit einem Zug-Icon installiert. In der vorliegenden Masterarbeit wird geprüft, ob die Maßnahmen die Wahrscheinlichkeit für einen adäquaten Kontrollblick im Vergleich mit der Kontrollbedingung signifikant erhöhen, wobei effektivere Ergebnisse seitens des Blinklichts erwartet werden. Außerdem wird anhand von Kennwerten der Interrater-Reliabilität geprüft, ob die Untersuchungsmethode objektive Ergebnisse liefert. Auf Basis von Videoaufnahmen wurden dazu bei 240 Querenden Kopfbewegungen nach links und rechts als Indikator für einen adäquaten Kontrollblick erfasst. In der Testbedingung mit der Markierung lag der Anteil der Personen, die adäquat schauten, mit 62 % signifikant höher als in der Kontrollbedingung (38 %). In der Testbedingung mit dem Blinklicht stieg der Anteil im Vergleich zur Bedingung mit der Markierung signifikant auf 76 %. Die Objektivität der Methode wurde durch den Vergleich mit einer unabhängigen Annotation bestätigt (Yules $Y = .614$). Der beobachtete Anteil von Personen mit adäquatem Kontrollblick blieb über die Testzeiträume konstant. Insgesamt zeigt die Masterarbeit, dass zusätzliche Maßnahmen zur Förderung der Sicherheit an Bahnübergängen empfehlenswert sind. Optimierungsideen bezüglich der Maßnahmen zielen zum einen auf die Erhöhung der Salienz der Markierung sowie auf die bessere Verständlichkeit der Maßnahmen bei einer breiteren Zielgruppe ab. In zukünftigen Untersuchungen sollte die Maßnahmenwirksamkeit über längere Zeit analysiert werden. Zusätzlich könnte eruiert werden, ob ein positiver Effekt auf das Kontrollblickverhalten auch bei anderen Straßenverkehrsteilnehmenden wie Autofahrenden auftritt.

Abstract

Safety at level crossings depends on the behaviour of road users, as they must give priority to rail traffic. The effects of accidents are serious, especially for vulnerable road users such as pedestrians or cyclists: in addition to the personal impact of fatalities and injuries, they cause considerable costs in terms of property damage, environmental damage and consequential delays. One focus of the SAFER-LC research project was to promote safe visual scanning behaviour before crossing among this target group. Within this framework, two innovative, low-cost measures were tested at a passive level crossing. From the development phase to the post-hoc evaluation of the measures, attention was paid to the fact that safety-relevant behaviours are promoted through environmental design. After a control survey, the measures were implemented fortnightly. On the one hand, a road marking with the words "Is a train coming?" was applied and on the other hand, an amber light with a train icon was installed. This master's thesis examines whether the measures significantly increase the probability of an adequate visual scanning behaviour in comparison with the control condition, whereby more effective results are expected on the part of the amber light. In addition, interrater reliability parameters are used to examine whether the test method provides objective results. On the basis of video recordings, head movements to the left and right were logged for 240 crossers as an indicator of adequate visual scanning behaviour. During the test phase with the marking, the proportion of people who looked adequately was significantly higher (62%) than in the control phase (38%). In the condition phase with the flashing light, the proportion rose again significantly to 76% compared to the condition with the marker. The objectivity of the method was confirmed by comparison with an independent annotation (Yules $Y = .614$). The observed proportion of subjects with adequate visual scanning behaviour remained constant over the test periods. In general, the master thesis shows that additional measures to promote safety at level crossings are recommendable. Ideas for optimising the measures were developed, which aim to increase the salience of the markings as well as could improve the comprehensibility of the two measures for a broader target group. In future studies, the effectiveness of the measures should be analysed over a longer period of time, to evaluate their long-term effectiveness or whether further measures improvement is needed. In addition, it could be investigated whether a positive effect on the visual scanning behaviour also occurs with other road users such as car drivers.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis.....	III
Tabellenverzeichnis	IV
Abkürzungsverzeichnis.....	V
1. Einleitung.....	1
1.1. Motivation und Zielsetzung.....	1
1.2. Das Projekt SAFER-LC.....	2
2. Theoretischer Hintergrund.....	4
2.1. Verhalten im Straßenverkehr als Reiz-Reaktions-Verknüpfung.....	4
2.2. Klassifikation und Ursachen von Fehlverhalten am BÜ	7
2.2.1. Fehlerklassifikation von Reason (1990).	8
2.2.2. Menschliche Faktoren bei Unfällen am BÜ nach Rudin-Brown, George und Stuart (2014).	10
2.2.3. Communications-Human Information Processing-Modell von Conzola und Wogalter (2001).	12
2.2.4. Protection Motivation Theory (PMT) von Rogers, Cacioppo und Petty (1983).	14
2.2.5. Anknüpfungspunkte für Interventionen.	17
2.3. Getestete Maßnahmen	18
2.3.1. Beschreibung der Maßnahmen.....	18
2.3.2. Interventionspsychologische Einordnung.....	21
2.4. Bewertung der Maßnahmen anhand der theoretischen Herleitung.....	22
2.5. Gewöhnungseffekt.....	24
2.6. Auswahl und Objektivität der Datenerhebungsmethode	26
2.7. Hypothesen	27
3. Methoden	30
3.1. Beschreibung des Settings	30
3.2. Strichprobengewinnung und -beschreibung	31
3.3. Ablauf der beiden Annotationen.....	32
3.4. Datenerhebungsmethode.....	33

3.5.	Statistische Auswertung.....	34
3.5.1.	Vergleich des Kontrollblickverhaltens der drei Bedingungen.....	35
3.5.2.	Auftreten des Gewöhnungseffekts.....	36
3.5.3.	Bestimmung der Interrater-Reliabilität.....	37
4.	Ergebnisse.....	38
4.1.	Maßnahmeneffektivität.....	38
4.2.	Gewöhnungs- und Wochentageffekt	40
4.3.	Interrater-Reliabilität	41
5.	Diskussion.....	44
5.1.	Maßnahmeneffektivität.....	44
5.2.	Gewöhnungs- und Wochentageffekt	46
5.3.	Interrater-Reliabilität	47
5.4.	Implikationen für die Praxis	48
5.5.	Limitationen der Untersuchungsmethode.....	50
5.6.	Offene Forschungsfragen.....	52
6.	Literaturverzeichnis	56
Anhang.....		
Anhang A Häufigkeitstabelle des differenzierten Kontrollblickverhaltens		A1
Anhang B Pearson Chi-Quadrat-Test zur Maßnahmeneffektivität		A2
Anhang C Häufigkeitstabelle zum adäquaten Kontrollblickverhalten		A4
Anhang D Pearson Chi-Quadrat-Test zum Gewöhnungseffekt		A5
Anhang E Pearson Chi-Quadrat-Test zum Wochentageffekt.....		A6
Anhang F Interrater-Reliabilitäten zu den Eigenschaften der VRU.....		A7
Anhang G Interrater-Reliabilitäten zum Kontrollblickverhalten der VRU.....		A9

Abbildungsverzeichnis

Abbildung 1	Drei-Ebenen-Modell nach Rasmussen (1983).....	5
Abbildung 2	Fehlerklassifikation von Reason (1990).	8
Abbildung 3	Communication-Human Information Processing Modell von Conzola und Wogalter (2001).	13
Abbildung 4	Protection Motivation Theory (Rogers et al., 1983).....	15
Abbildung 5	Erste getestete Maßnahme: Straßenmarkierung.....	19
Abbildung 6	Zweite getestete Maßnahme: Blinklicht.	20
Abbildung 7	Beurteilung der Maßnahmen anhand der von Conzola und Wogalter (2001) beschriebenen Attribute für die Gestaltung effektiver Warnhinweise sowie anhand der im PMT von Rogers et al. (1983) beschriebenen Kernkomponenten.....	23
Abbildung 8	Bildaufnahmen des Settings.....	30
Abbildung 9	Stichprobengewinnung innerhalb des Versuchsablaufs.....	31
Abbildung 10	Einteilung der Messzone anhand einer Momentaufnahme der Videodaten.	34
Abbildung 11	Differenzierte Darstellung des Kontrollblickverhaltens innerhalb der Testbedingungen.	38
Abbildung 12	Prozentuales Auftreten adäquaten Kontrollblickverhaltens an den acht Messtagen der Testbedingungen mit Maßnahmen.	40
Abbildung 13	Zwei Ideen zur Optimierung des Blinklichts zur Vermeidung von Fehlinterpretationen.	49
Abbildung 14	Idee zur Optimierung der Straßenmarkierung zur Erhöhung der Salienz.....	50
Abbildung 15	Idee zur Optimierung der Straßenmarkierung zur Verbesserung der Verständlichkeit auch für Personen, die (noch) nicht lesen können.....	50

Tabellenverzeichnis

Tabelle 1	Häufigkeiten des adäquaten Kontrollblickverhaltens innerhalb der Testbedingungen	39
Tabelle 2	Analyse der Interrater-Reliabilität zu den Eigenschaften der VRU.....	42
Tabelle 3	Analyse der Interrater-Reliabilität zu dem beobachteten Kontrollblickverhalten der VRU	43
Tabelle 4	Häufigkeiten des differenzierten Kontrollblickverhaltens innerhalb der Bedingungen.....	A1
Tabelle 5	Häufigkeitstabelle der Analysen zur Maßnahmeneffektivität.....	A2
Tabelle 6	Ergebnisse der Chi-Quadrat-Tests zur Maßnahmeneffektivität.....	A3
Tabelle 7	Häufigkeitstabelle zum adäquaten Kontrollblickverhalten der acht Messtage der Testbedingungen mit Maßnahmen.....	A4
Tabelle 8	Häufigkeitstabellen der Analysen zum Gewöhnungseffekt der beiden Testphasen mit Maßnahmen.....	A5
Tabelle 9	Ergebnisse der Chi-Quadrat-Tests zu den Gewöhnungseffekten.....	A5
Tabelle 10	Häufigkeitstabellen der Analysen zum Wochentageffekt der beiden Testphasen mit Maßnahmen.....	A6
Tabelle 11	Ergebnisse der Chi-Quadrat-Tests zu den Wochentageffekten.....	A6
Tabelle 12	Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich des Geschlechts der VRU	A7
Tabelle 13	Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich der Altersgruppe der VRU.....	A7
Tabelle 14	Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich der Fortbewegungsart der VRU.....	A8
Tabelle 15	Häufigkeitstabelle zur Übereinstimmung der Beurteilung adäquaten Kontrollblickverhaltens zwischen Annotation 1 und 2.....	A9
Tabelle 16	Häufigkeitstabelle zur Übereinstimmung der Beurteilung des differenzierten Kontrollblickverhaltens zwischen Annotation 1 und 2	A9

Abkürzungsverzeichnis

AIM	Anwendungsplattform Intelligente Mobilität (Forschungsanlage des DLR; Kombination verschiedener Analyseverfahren im Baukastenprinzip möglich)
BÜ	Bahnübergang
C-HIP	Communications-Human Information Processing (Conzola & Wogalter, 2001)
DLR	Deutsches Zentrum für Luft- und Raumfahrt e.V.
PBC	Perceived Behavior Control
PMT	Protection Motivation Theory (Rogers, Cacioppo & Petty, 1983)
SAFER-LC	Safer Level Crossing by integrating and optimizing road-rail infrastructure management and design (übergeordnetes Projekt)
VRU	vulnerable road users (ungeschützte Straßennutzende; zu Fuß Gehende, Radfahrende etc., die keine protektiv wirkende Pufferzone bei Unfällen um sich haben)

1. Einleitung

1.1. Motivation und Zielsetzung

In Deutschland existieren 16.541 Bahnübergänge (BÜ) der Deutschen Bahn AG (Deutsche Bahn AG, 2019) sowie etwa die gleiche Anzahl an BÜ von Privatbahnen (Dressler & Grippenkov, 2018). Die Hauptunfallursache an BÜ war 2017 die Missachtung des Vorrangs der Schienenfahrzeuge durch Straßenverkehrsteilnehmende, die nach Angaben der Deutschen Bahn AG durch willentliche Verstöße oder Unachtsamkeit geschah (Deutsche Bahn AG, 2019). Insgesamt starben deutschlandweit 35 Personen bei der Benutzung eines BÜ, was einen Anteil von rund 27 % aller Verstorbenen in Folge von Eisenbahnunfällen ausmacht (Eisenbahn-Bundesamt, 2018). Die Auswirkungen von Unfällen sind auf Grund der physischen Eigenschaften von Zügen schwerwiegend: Sie verursachen neben den persönlichen und wirtschaftlichen Auswirkungen durch die Todesfälle und Verletzungen erhebliche Kosten durch Sachschäden, Umweltschäden und Folgeverspätungen aufgrund von Unterbrechungen des Eisenbahnbetriebs (European Union Agency for Railways, 2018). Aufgrund der Kosten und infrastrukturellen Gegebenheiten „ist es kaum möglich, alle Bahnübergänge zu schließen oder mit bewährten technischen Sicherungssystemen wie Schranken und Lichtsignalanlagen aufzuwerten“ (Dressler & Grippenkov, 2018, S. 26). Generell verfolgen Verkehrsunternehmen den Ansatz, höhengleiche BÜ, an denen sich Eisenbahn- und Straßenverkehr auf einer Ebene begegnen, zu vermeiden. So hat sich die Zahl der BÜ seit 1950 mehr als halbiert, die Tendenz ist weiter sinkend (Schumacher & Weihgold, 2017). Dies erreichen sie durch ein Verschließen der Querungsmöglichkeit und durch Tunnel- oder Brückenbau (Dressler & Grippenkov, 2018).

Ein besonders hohes Gefährdungspotenzial bieten nicht-technisch gesicherte BÜ. Der Anteil nicht-technisch gesicherter BÜ der Deutschen Bahn AG macht rund 38 % aller BÜ aus (DB Netze, 2018). Der Anteil aller Unfälle zwischen Eisenbahn- und Straßenverkehr an nicht-technisch gesicherten BÜ lag 2018 bei 53 % (DB Netze, 2018), dabei entstanden 76 Personenschäden mit zwölf Schwerverletzten und zehn Todesfällen. Sicheres Querungsverhalten liegt maßgeblich in der Verantwortung der Straßenverkehrsteilnehmenden. Das ist durch das Bundesministerium der Justiz und für Verbraucherschutz in der Straßenverkehrsordnung (StVO) in entsprechenden Paragraphen verankert: In § 19 der StVO, Absatz 1 und 2 zu Bahnübergängen ist festgelegt, dass Schienenfahrzeuge Vorrang haben, sich der Straßenverkehr solchen Bahnübergängen „nur mit mäßiger Geschwindigkeit nähern“ darf sowie, dass Fahrzeugfahrende „vor dem Andreaskreuz, zu Fuß Gehende in sicherer Entfernung vor dem

Bahnübergang“ warten müssen. Das Andreaskreuz verdeutlicht als Warte- und Haltegebots-schild diese Vorschrift am BÜ: „Wer ein Fahrzeug führt, muss dem Schienenverkehr Vorrang gewähren“ (Bundesministerium der Justiz und für Verbraucherschutz, 2020). Über die Gesetz-lage hinaus verhindert die „Schienenbindung ... ein Ausweichen“ und „die hohe Ge-schwindigkeit der Züge gepaart mit dem hohen Gewicht und der geringen Haftreibung zwi-schen Stahlrad und Stahlschiene machen es dem Triebfahrzeugführer selbst bei einer frühzei-tigen Bremsreaktion kaum möglich, den Zug noch vor einem Hindernis zum Stillstand zu bringen“ (Grippenkoven, Thomas & Lemmer, 2016, S. 48). Besonders ernsthafte Folgen von Eisenbahnunfällen sind bei ungeschützten Straßennutzenden (engl. *vulnerable road-user*, VRU) wie zu Fuß Gehende oder Fahrradfahrende zu erwarten, da sie keine protektiv wirkende Pufferzone wie Autofahrende um sich haben. Diese Probleme verdeutlichen „the need to take a human-centered approach that considers perception and attention as well as judgment and motivation“ (Dressler et al., 2020, S. 2) zur Vermeidung von Unfällen zwischen Schienen-fahrzeugen und VRU in diesem Setting. Das zentrale Ziel der vorliegenden Masterarbeit ist, die Wirksamkeit zweier getesteter Sicherungsmaßnahmen an nicht-technisch gesicherten BÜ auf das Verhalten bei der Querung von VRU zu evaluieren. Darüber hinaus wird exploriert, ob sich innerhalb der Testphasen bereits die zu erwartenden Gewöhnungseffekte an derartige Interventionen eingestellt haben könnten.

Für die empirische Untersuchung von Phänomenen menschlichen Erlebens und Ver-haltens wurden die klassischen Gütekriterien der Objektivität, Reliabilität und Validität ent-wickelt, die zur nachträglichen Beurteilung der wissenschaftlichen Qualität herangezogen werden können und sollten (Hussy, Echterhoff & Schreier, 2010). Daher ist über die Wirk-samkeitsbewertung der Maßnahmen hinaus eine Beurteilung der Objektivität der Datenerhe-bungsmethode in der vorliegenden Masterarbeit durchgeführt worden. Eine erste Auswertung der Maßnahmenwirksamkeit hat bereits im Rahmen einer Abschlussarbeit stattgefunden (Dressler, Illgen & Grippenkoven, 2020; Illgen, 2020). Auf Basis der Annotationen von Illgen (2020) wurde die Interrater-Reliabilität als Maß für die Objektivität der gewählten Untersu-chungs- und Auswertungsmethode mit dem verfügbaren Equipment herangezogen.

1.2. Das Projekt SAFER-LC

Im Rahmen des europäischen Verbundprojektes SAFER-LC (*Safer level crossing by integra-ting and optimizing road-rail infrastructure management and design*) wurden zwischen Mai 2017 und April 2020 Ansätze zur Verbesserung der Straßen- und Schienensicherheit entwi-kelt, wobei sowohl technische Lösungen als auch menschliche Prozesse betrachtet wurden

(Dressler et al., 2020). Die Koordination und Projektleitung übernahm der Internationale Eisenbahnverband UIC. An dem Projekt waren nach Dressler und Gripenkoven (2018) 17 Verbundpartner aus zehn Ländern beteiligt. Dazu gehörten Bahn- und Schienennetzbetreibende, Verkehrsforschungszentren, Technische Hochschulen, Technologieunternehmen für Verkehrs-Infrastruktur-Kommunikation und Vertreter:innen der Straßenverkehrsträger. Diverse Expert:innen aus der internationalen Eisenbahn- und Straßenverkehrspraxis sowie Vertreter:innen von Verkehrsministerien und Organisationen für Sicherheit im Straßenverkehr unterstützten zusätzlich. Dadurch, dass Verbundpartner verschiedener Fachrichtungen an dem Projekt SAFER-LC arbeiteten, entstand eine systemische Analyse und Maßnahmenentwicklung für die Förderung der Straßen- und Schienensicherheit Europas (Dressler et al., 2020). Ein Fokus des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) lag auf der Zielgruppe der VRU wie zu Fuß Gehende und Fahrradfahrende, die bei Eisenbahnunfällen an BÜ oft noch schwerwiegendere Verletzungen als andere Straßenverkehrsteilnehmende erleiden, da eine protektiv wirkende Pufferzone fehlt. Die infrastrukturellen Gegebenheiten, die ein bestimmtes Verhalten der VRU begünstigen, spielen eine bedeutsame Rolle bei der Gestaltung von sicheren BÜ: Wahrnehmung, Aufmerksamkeits- und Beurteilungsprozesse und die Motivation seitens der VRU sollten bei der Maßnahmengestaltung berücksichtigt werden, wenn durch die Umweltgestaltung sicherheitsrelevante Verhaltensweisen gefördert werden sollen (Dressler et al., 2020). Das DLR entwickelte und testete kostengünstige Gestaltungsmöglichkeiten zur Erhöhung der Auftretenswahrscheinlichkeit von adäquatem Kontrollblickverhalten an nicht-technisch gesicherten BÜ. Kostengünstige Maßnahmen wurden nach Dressler et al. (2020) definiert als Maßnahmen, die günstiger sind als konventionelle Optimierungen wie beispielsweise Lichtzeichenanlagen oder Schranken. Potenzielle Maßnahmen für einen experimentellen Testbetrieb sollten außerdem einen Innovationscharakter aufweisen und derzeit europaweit noch unüblich sein. Adäquates Kontrollblickverhalten wird in der vorliegenden Masterarbeit definiert als Blick nach rechts und links nach einem Zug. Inadäquat ist dementsprechend, wenn die Querenden gar nicht oder nur in eine Richtung nach einem Schienenfahrzeug Ausschau halten. Die drastische Kategorisierung, dass der Kontrollblick in nur eine Richtung und in gar keine Richtung gleichgesetzt werden, beruht auf der Annahme, dass das Verhalten im Straßenverkehr per se entweder richtig sein kann und somit wahrscheinlich zu keinem Unfall führt, oder falsch sein kann und die Wahrscheinlichkeit für einen Unfall erhöht ist. Wenn die Person nur in eine Richtung schaut, hat die Maßnahme nicht zur angestrebten Vermeidung unsicherer Verhaltensweisen geführt.

2. Theoretischer Hintergrund

Die Verkehrspsychologie „untersucht die Interaktionsbeziehungen zwischen Verkehrssystemen und menschlichem Erleben und Verhalten“ (Richter, 2016, S. 48). Das Ziel der Verkehrspsychologie ist, diese Wechselwirkungsbeziehung zwischen dem Menschen und dem Verkehrssystem derart zu optimieren, dass gewünschtes Verhalten von den strukturellen Gegebenheiten gefördert wird. Das menschliche Verhalten im Straßenverkehr lässt sich über eine Vielzahl psychologischer Modelle zur kognitiven Informationsverarbeitung beschreiben. Grundsätzlich operieren kognitive Verhaltensmodelle nach Städtler „mit der zentralen Annahme, dass der menschliche Geist einem Informationsverarbeitungsmodell gleicht, wie es ... in einem Computer realisiert ist“ (1998, S. 479). Weiterhin wird der Informationsverarbeitungsprozess hierarchisch in drei Stufen unterteilt: Aufnahme, Informationsspeicherung, Abruf. Unterschieden werden grundsätzlich die kontrollierte Informationsverarbeitung mit Aufmerksamkeit, die selektive Aufmerksamkeitsprozesse bedingt, und die automatischen Prozesse ohne Aufmerksamkeit, die keine mentale Anspannung (engl. *effort*) benötigen, wobei der Übergang zwischen diesen beiden oftmals, z.B. von Rasmussen (1983), als Kontinuum verstanden wird.

Zwei mögliche Maßnahmen für Förderung des sicheren Verhaltens bei der Querung wurden im Rahmen des Projektes SAFER-LC entworfen und getestet. In der vorliegenden Masterarbeit werden diese Maßnahmen auf der Grundlage von aktuellen Studienbefunden zunächst eingeordnet. Hierfür werden zunächst Überlegungen zu dem Verhalten am BÜ auf Basis gängiger kognitiver Verhaltensmodelle angestellt. Anschließend werden mögliche Fehler im Informationsverarbeitungsprozess dargestellt, woraufhin Theorien zur Gestaltung innovativer und kostengünstiger Maßnahmen zur Vermeidung ebendieser vorgestellt werden. Auf dieser Basis können schließlich Hypothesen für die Datenauswertung generiert werden. Ein grundsätzliches Problem dieser Art von Interventionen, die in der Studie erprobt wurden, sind Gewöhnungseffekte. Sie werden die in der vorliegenden Masterarbeit ebenfalls beschrieben und Befunde dazu vorgestellt, sodass eine Hypothese zu dessen Auftreten formuliert werden kann. Abschließend werden noch die Themen Datenerhebungsmethode und Objektivität beleuchtet, um die Relevanz der Interrater-Reliabilität für die Generalisierbarkeit der Befunde sowie für die Vermeidung von Verzerrungen zu verdeutlichen.

2.1. Verhalten im Straßenverkehr als Reiz-Reaktions-Verknüpfung

Bei den Forschungen in der Verkehrspsychologie und im Bereich Human Factors wird nach Groeger (2000) grundsätzlich davon ausgegangen, dass das Verhalten im Straßenverkehr

Rasmussen (1983) unterscheidet hierbei (a) das wissensbasierte Verhalten (*knowledge based*), das hochgradig bewusst abläuft. Diese Art von Verhalten beinhaltet komplexe Problemlösung, da weder Routinen noch Regeln für die erfolgreiche Situationsbewältigung vorhanden sind sowie als Improvisationen in unbekanntem Umgebungen beschrieben werden können. (b) Das regelbasierte Verhalten (*rule based*), bei dem bereits Reiz-Reaktions-Verknüpfungen im Sinne von Assoziationen zwischen Umgebungszuständen, zu lösender Probleme und daraus resultierender Aufgaben für die Person existieren. In der von Städtler (1998) beschriebenen kognitiven Verarbeitung des Menschen entsprechend der Informationsverarbeitung bei Computern basiert diese Art des Verhaltens auf *IF... THEN...* Schleifen. Schließlich folgt (c) das auf sensomotorischen Fertigkeiten basierte Verhalten (*skill based*), welches hochgradig automatisiert abläuft, kaum einer Aufmerksamkeitszuwendung bedarf und als Routine bezeichnet werden kann.

Entsprechend der drei Ebenen liegt das Verhalten im Straßenverkehr im regelbasierten Bereich, wobei an dieser Stelle der Grad der Erfahrung im Straßenverkehr generell sowie auch in der jeweiligen Umgebung betont werden muss, die den Anteil von zusätzlichen bewussten Anteilen im kognitiven Prozess bedingen. Nach Reason (1990) wird ein Verhalten umso automatischer, je häufiger die Person den entsprechenden Reiz präsentiert bekommt und die Reaktion zu einem positiven Ergebnis führt. Hierbei wird die Verknüpfung zwischen Reiz und Reaktion so weit gestärkt, bis der Anteil bewusster Komponenten so niedrig wie möglich wird, jedoch selten vollständig verschwindet. Ein klassisches regelbasiertes Verhalten mit automatischen und bewussten Anteilen ist entstanden. Sollte die Reaktion nicht mehr zu einem gewünschten Ergebnis führen, wird ein arbeitsaufwändiger Modus zur Korrektur der Regel aktiviert. Aus dem vorhandenen Wissen wird wissensbasiert nach Lösungsstrategien gesucht, um die Situation zu bewältigen. Zeigen sich wiederum permanente Erfolge durch die neu entwickelte Verhaltensstrategie auf den Reiz, führt dies zu einer neuen starken Reiz-Reaktions-Verknüpfung.

Durch die Einordnung als regelbasiertes Verhalten wird deutlich, dass das Verstehen des Umweltzustands „ein konstruktiver Prozess ist, an dem außer der direkt gegebenen Information weitere Information – unser Wissen – beteiligt ist“ (Schwarz, 1985, S. 1). Mentale Modelle und kognitive Schemata spielen hierbei eine Rolle. Mentale Modelle sind nach Städtler (1998) geistige Abbildungen eines Wirklichkeitsbereichs und können entsprechend ihres Detaillierungsgrades zwischen propositionalen Wissensrepräsentationen und einem konkreten Vorstellungsbild eingeordnet werden. Kognitive Schemata sind nach Städtler

(1998, S. 951) „langfristig gespeicherte, hierarchisch gegliederte Wissensstrukturen“, in denen z.B. Umweltkonstellationen und Handlungsfolgen mental abgebildet werden. Aktuell wahrgenommene Informationen werden top-down in existierende Schemata im Sinne von Schablonen eingeordnet und dadurch erkannt und verstanden. Bezogen auf den BÜ sind mentale Modelle die Summe der subjektiven Annahmen der Person, wie ein BÜ funktioniert, u.a. auf Grund früherer Erfahrungen, der aktuellen Wahrnehmung und bereits angelegter Problemlösestrategien. Kognitive Schemata integrieren alle Informationsfragmente in die Umweltkonstellation *Bahnübergang*, woraufhin idealerweise eine adäquate Handlungsfolge (Skript) aktiviert wird, die das Kontrollblickverhalten und, sofern anwendbar, das Abbremsen des eigenen Fahrzeugs beinhaltet. Durch diese Schematisierung kann der Mensch nach Esser (1990) beispielsweise schnell und adäquat in ihm bekannten Situationen reagieren. „Es handelt sich dabei um allgemeine und integrierte Wissensstrukturen über einen bestimmten Bereich, die den [Prozeß] der Informationsverarbeitung drastisch vereinfachen und so dem Akteur eine relativ leichte Situations-Orientierung erlauben“ (Esser, 1990, S. 234). Zachrisson und Boks (2010) nehmen bezüglich der Veränderung von automatischem, habituiertem Verhalten an, dass der Person das Verhalten bewusst gemacht werden sollte, wenn deren Einstellung geändert oder neue, nachhaltige Gewohnheiten geschaffen werden sollen. Es ist entscheidend, dass die Ausführung einer ungewünschten Gewohnheit unterbrochen wird, um mit neuen intentionalen oder situativen Faktoren zu einem positiveren Verhalten und einer neuen Gewohnheit zu führen. Die Situation am BÜ kann als Interaktion zwischen dem Menschen und der Infrastruktur betrachtet werden, bei der es nach Zachrisson und Boks (2010) für eine erfolgreiche Veränderung des automatischen Verhaltens hilft, wenn die implementierten Maßnahmen möglichst innovativ sind und die Interaktion dadurch möglichst andersartig gestaltet wird.

2.2. Klassifikation und Ursachen von Fehlverhalten am BÜ

Die Bedeutsamkeit sicheren Verhaltens am BÜ seitens der Straßenverkehrsteilnehmenden wurde in der Einleitung verdeutlicht. Im Folgenden werden Modelle und Theorien dazu vorgestellt, welche Fehlverhaltensweisen auftreten können und welchen Ursprung sie an einem nicht-technisch gesicherten BÜ haben können.

2.2.1. Fehlerklassifikation von Reason (1990).

Auf den Überlegungen zu den beiden Extrema bewusst vs. automatisch sowie den drei Ebenen klassifizierte der Fehlerforscher Reason (1990) entsprechende Fehlerarten, die auftreten können und in Abbildung 2 dargestellt sind. Sie werden im Folgenden mit Beispielen zu möglichem Fehlverhalten bei der Querung eines BÜ verdeutlicht.

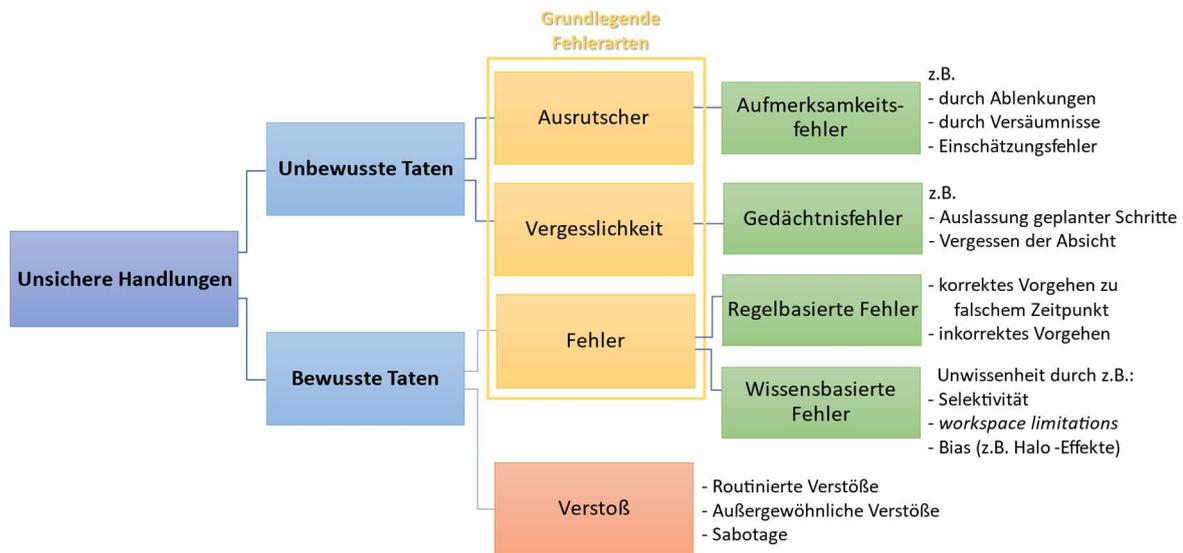


Abbildung 2 Fehlerklassifikation von Reason (in Anlehnung an Reason, 1990, S. 207).

Reason (1990) unterteilt unsichere Handlungen in unbewusste und bewusste Taten. Unbewusste Taten unterteilen sich wiederum in zwei Kategorien: (a) Ausrutscher, die auf Aufmerksamkeitsfehlern beruhen. Dazu gehören Fehler, die auf Grund von Ablenkungen (Person übersieht Andreaskreuz an BÜ, da sie nebenbei eine Nachricht auf dem Smartphone tippt), Versäumnisse (Person bremst Fahrrad an BÜ nicht ab, da sie telefoniert und zu spät daran denkt) sowie Einschätzungsfehlern (Person schätzt Weg zum BÜ kürzer ein und hat daher nur noch Zeit, in eine Richtung nach einem Schienenfahrzeug zu schauen) entstehen. (b) Vergesslichkeit, die auf Grund von Gedächtnisfehlern entsteht. Hierbei können geplante Schritte ausgelassen werden (Person vergisst während des Schauens nach einem Schienenfahrzeug abzubremsen und befindet sich schneller als erwartet auf dem BÜ) oder die Absicht kann vergessen werden (Person vergisst in eine Richtung zu schauen, da sie beim Blick in die andere Richtung eine interessante Entdeckung gemacht hat). Unsichere Handlungen können jedoch auch durch bewusste Taten entstehen. Die bewussten Taten werden unterteilt in zwei Kategorien: (c) Klassische Fehler, die zwar bewusst, aber nicht absichtlich geschehen. Sie können zum einen durch regelbasierte Fehler entstehen. Hierzu gehören Fehler, bei denen ein

korrektes Vorgehen zur falschen Zeit durchgeführt wird. Dies geschieht, wenn eine Reiz-Reaktionsverknüpfung stark ist und in einer neuen Umgebung zwar ähnliche Reize vorliegen, jedoch andere Reaktionen erwartet werden (Person kennt nicht-technisch gesicherte BÜ aus ihrem Alltag nicht. Daher erwartet sie nicht, dass dort Schienenfahrzeuge kreuzen, ohne dass dies durch eine technische Sicherung angezeigt wird. Sie quert ohne adäquates Verhalten). Eine weitere Fehlerart des klassischen Fehlers ist das inkorrekte Vorgehen auf Grund falscher Reiz-Reaktions-Verknüpfungen (Person zeigt kein adäquates Verhalten am BÜ, da sie hier bisher noch nie ein Schienenfahrzeug kreuzen sah; Es hat sich an diesem BÜ ein „Kein-Zug-Schema“ etabliert). Zu den klassischen Fehlern gehören ebenso die wissensbasierten Fehler, die durch Unwissenheit entstehen. Ursachen dafür sind z.B. Selektivität (Person achtet aus Sicherheitsgründen auf das auffällige Fahrverhalten des Fahrradfahrenden vor ihr und übersieht den nahenden BÜ), workspace limitations (Person übersieht bei Platzregen auf der vollen Straße den BÜ auf Grund von Reizüberforderung) oder Verzerrungen beispielsweise durch den Halo-Effekt (Person hat das Vorurteil, dass nicht-technisch gesicherte BÜ nicht besonders hoch frequentiert sind und vernachlässigt den ausgiebigen Kontrollblick). Die weitere Kategorie der bewussten Taten bezeichnet Reason (1990) nicht als klassische Fehler: (d) Verstöße geschehen bewusst und im Gegensatz zu Fehlern absichtlich. Er unterscheidet drei Arten von Verstößen: entweder solche, die routiniert geschehen (Person fährt willentlich immer rasant und bremst am BÜ nicht ab), die außergewöhnlich passieren (Person hat es eilig und fährt daher ohne abzubremesen) sowie ferner die Sabotage, bei der willentlich das System gefährdet wird (Person demoliert das Andreaskreuz und gefährdet dadurch andere).

Das Fazit aus den Überlegungen zu den möglichen Fehlern, die am BÜ geschehen können basierend auf der Klassifikation von Reason (1990), kann festgehalten werden, dass es sinnvoll ist, bei der Maßnahmengestaltung sowohl bewusste als auch unbewusste Handlungen zu betrachten. Die Maßnahmen sollten auch vor dem Begehen von bewussten Verstößen dazu motivieren, sicheres Verhalten bei der Querung zu zeigen, um sich und andere zu schützen. Ausrutscher aufgrund von kleineren Aufmerksamkeitsfehlern (z.B. interne Ablenkung durch Grübeln oder externe Ablenkung durch ein Gespräch), das Vergessen von Aspekten des sicheren Verhaltens am BÜ (z.B. Person schaut nur in eine Richtung) sowie regelbasierte Fehler (z.B. durch die Existenz eines „Kein-Zug-Schemas“) sollte mit Interventionen entgegen gewirkt werden, die die Aufmerksamkeit und den Fokus auf die nahende Situation und dessen Reflektion im Sinne einer bewussten Auseinandersetzung lenken.

2.2.2. Menschliche Faktoren bei Unfällen am BÜ nach Rudin-Brown, George und Stuart (2014).

Um das Unfallrisiko an nicht-technischen BÜ zu verringern, sollten die genauen Faktoren analysiert werden, die zu adäquatem Sicherheitsverhalten führen. Rudin-Brown, George und Stuart untersuchten 2014 die menschlichen Faktoren, die bei der Querung und bei Unfällen an BÜ eine Rolle spielen. Sie nutzten hierfür die Kanadischen Unfallzahlen von BÜ für den Zeitraum zwischen 2003 und 2012 in Kombination mit einer gezielten Überprüfung der wissenschaftlichen Forschungsliteratur. Grundsätzlich unterscheiden sie zwei essenzielle Kategorien von Faktoren, auf die die Unfallursachen rückführbar waren: Vier perzeptive Faktoren, die die Erkennung und Wahrnehmung des nicht-technisch gesicherten BÜ oder nahende Schienenfahrzeuge bedingen und fünf kognitive Faktoren, die die Entscheidungsfindung des Fahrzeugfahrenden beeinflussen, adäquat zu reagieren. Als Faktoren, die die Perzeption von BÜ beeinflussen, werden (a) behinderte Sichtlinien evtl. in Kombination mit den Eigenschaften des eigenen Autos (beispielsweise der Anfahrtswinkel zum BÜ in Kombination mit Fahrzeugsäulen genau im Blickfeld zur Bahnstrecke), (b) die Sichtbarkeit des Zuges (kritisch in der Dämmerung oder nachts), (c) das unveränderte Netzhautbild eines herannahenden Zuges (erschwert die Einschätzung der Geschwindigkeit) und (d) Schwierigkeiten bei der akustischen Wahrnehmung des Zughorns (gilt an nicht-technisch gesicherten BÜ, an denen der Schienenfahrzeugführende das Signalhorn ertönen lassen muss) genannt. Die Faktoren im Zusammenhang mit der Entscheidungsfindung des Autofahrenden sind (a) der so genannte "looked-but-failed-to-see Fehler" (eine Situation, in der es aufgrund von Einschränkungen in der menschlichen Informationsverarbeitung nicht gelingt, Gefahren zu erkennen, obwohl die Gefahrenquelle angesehen wird), (b) Ablenkungen (Aufmerksamkeit des Fahrenden entweder bei anderen auffälligen Umgebungsdetails oder bei internen kognitiven Denkprozessen), (c) die Beeinträchtigung des Triebfahrzeugführers (z.B. Alkohol oder Müdigkeit), (d) fehlerhafte kognitive Informationsverarbeitung sowie (e) die fehlerhafte Aktivierung korrekter Schemata und mentaler Modelle. Bei letzterem könnte beispielsweise ein „Kein-Zug-Schema“ gespeichert sein, das dann verstärkt und erlernt wird, wenn ein Straßenverkehrsteilnehmer mit einem bestimmten BÜ stark vertraut ist und in der Vergangenheit noch nie ein Zug während der Querung wahrgenommen wurde. Besonders fatal ist daran, dass eventuell regelmäßig ein adäquates Kontrollblickverhalten stattgefunden hat, sich jedoch durch die permanente Abwesenheit einer akuten Bedrohung die Gefährlichkeitseinschätzung der Situation verringert hat.

In einer Studie von Gripenkoven und Dietsch (2015) fuhren 24 Versuchspersonen eine definierte Strecke mit einem Forschungsfahrzeug, wobei Eyetracking-Daten und die Geschwindigkeit des Fahrzeugs aufgezeichnet wurden. Genauer untersucht wurden zum einen das Blickverhalten sowie die Geschwindigkeit bei der Annäherung vor der Querung eines nicht-technisch gesicherten BÜ. Durchschnittlich 15 % der visuellen Fixierungszeit wurde während der Annäherung auf die Bahnübergangsschilder verwendet, was als ausreichender Prozentsatz bewertet wird, damit ein Fahrzeugführender den nahenden BÜ perzeptiv wahrnimmt. Das Abbremsen des Fahrzeugs von erlaubten 50 km/h auf durchschnittlich 38.6 km/h fand im Bereich von 50-80 m vor dem BÜ statt, weshalb davon ausgegangen werden kann, dass alle Versuchspersonen den BÜ in angemessener Entfernung als solchen erkannten. Adäquates Verhalten bei der Querung bezüglich des Kontrollblicks und damit einhergehend die Reduzierung der Geschwindigkeit sind zeitlich möglich. In Bezug auf die von Rudin-Brown et al. (2014) beschriebenen menschlichen Faktoren, die zu adäquatem Sicherheitsverhalten führen, kann folglich der Schluss gezogen werden, dass das Problem des maladaptiven Verhaltens kein Problem der Perzeption ist. Anschließend zeigten jedoch nur 16 Versuchspersonen ein adäquates Kontrollblickverhalten an nicht-technisch gesicherten BÜ nach einem Schienenfahrzeug. Untersucht wurde außerdem, ob die Versuchspersonen, die den Kontrollblick adäquat durchführten, die Geschwindigkeit des Fahrzeugs stärker reduzierten. Hierbei konnte eine signifikante Reduzierung der Versuchspersonen mit Kontrollblick auf eine Geschwindigkeit von durchschnittlich 33.2 km/h festgestellt werden, wohingegen die Reduzierung bei den Versuchspersonen ohne Kontrollblick bei durchschnittlich 41.9 km/h lag. Dieser signifikante Unterschied sollte darauf zurückzuführen sein, dass ein adäquates Schauen entsprechend Zeit benötigt, wofür die Personen das Fahrzeug entsprechend bremsen. Gripenkoven und Dietsch (2015) haben für das Ausbleiben des Kontrollblicks zwei mögliche Erklärungsansätze. Zum einen schlussfolgern sie aus ihren Befunden, dass der BÜ nur in einem vorzeitigen sensorischen Stadium verarbeitet wurde, aber keine tiefere Verarbeitungsstadien wie die Identifizierung und das Auffinden der damit verbundenen Verhaltensschemata und -skripte aktiviert wurden. Gripenkoven und Dietsch (2015) vergleichen diese Art von Fehler mit dem von Rudin-Brown et al. (2014) beschriebenen „looked-but-failed-to-see“-Fehler, bei dem der nicht-technisch gesicherte BÜ an sich die Gefahr darstellt: Bei einem nicht-technisch gesicherten BÜ sollte allen Querenden bewusst sein, dass jederzeit ein Schienenfahrzeug kreuzen könnte. Durch diese Gedächtnis- oder Aufmerksamkeitsfehler entsprechend der Klassifikation durch Reason (1990) entstehen nicht nur unsichere Handlungen, sondern lebensgefährliche Situationen. Eine weitere andere Erklärung für das Fehlen des adäquaten

Verhaltens ist, dass der Großteil der Personen nicht weiß, wie sie sich an einem nicht-technisch BÜ verhalten sollen, weil sie weder über geeignete mentale Modelle oder Schemata noch über entsprechende Verhaltensskripte verfügen. Dies kann z.B. der Fall sein, wenn den Versuchspersonen eher technisch gesicherte BÜ geläufig sind und sie trotz Abwesenheit entsprechender Anzeigen für das Herannahen eines Schienenfahrzeugs ein Signal erwarten. Bezogen auf die von Reason (1990) genannten Fehlerarten entspricht dies den Fehlern aufgrund von maladapтивem regel- oder wissensbasiertem Verhalten.

Als Fazit der zusammengefassten menschlichen Faktoren von Rudin-Brown et al. (2014) und der Befunde der Studie von Grippenkoven und Dietsch (2015) sollten die Maßnahmen eine tiefere kognitive Verarbeitung über die Wahrnehmung und Identifikation hinaus anregen. Dies kann entweder bedeuten, dass vorhandene Verhaltensskripte aktiviert werden oder aber bei Nicht-Vorhandensein dazu führen, dass sich die Personen bewusst mit der nahenden Situation auseinandersetzen, eine entsprechend adäquate Verhaltensstrategie entwickeln und diese idealerweise langfristig als Verhaltensskript angelegt wird.

2.2.3. Communications-Human Information Processing-Modell von Conzola und Wogalter (2001).

Das erweiterte Communications-Human Information Processing (C-HIP)-Modell ist ein von Conzola und Wogalter (2001) entwickeltes Modell, um die Stufen bei der kognitiven Verarbeitung von Warnhinweisen durch die Mitarbeitenden in Unternehmen zu beschreiben. Es beinhaltet die relevanten Stufen einer (Warn-) Information von der Quelle oder einem Sendenden zu einem Empfangenden, der dann die Informationen verarbeitet, um anschließend ein Verhalten zu erzeugen. Das C-HIP-Modell wird im Folgenden genauer beschrieben und ist in Abbildung 3 dargestellt.

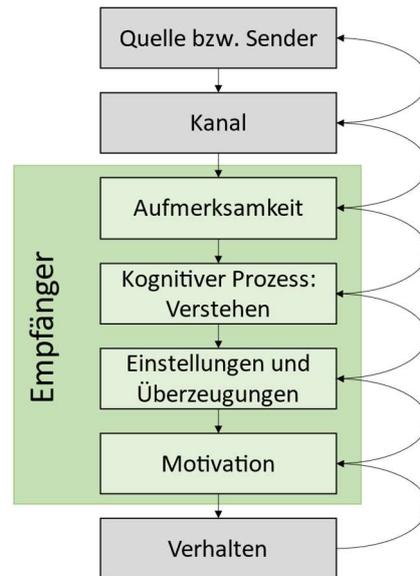


Abbildung 3 Communication-Human Information Processing (C-HIP-) Modell von Conzola und Wogalter (eigene Darstellung, 2001, S. 312).

Die Information geht von der Quelle oder von dem Sendenden aus über einen bestimmten Kanal zu dem Empfangenden. Die geistigen Aktivitäten des Empfangenden lassen sich in eine Abfolge von Informationsverarbeitungsstufen einordnen. Zunächst muss die Warnung die Aufmerksamkeit auf sich lenken und dann so lange aufrechterhalten, bis der Empfangende die notwendigen Informationen aus der Warnung extrahiert hat. Als nächstes muss die Warnung verstanden werden, damit Informationen zu einer potenziell gefährdenden Situation und gewünschtes Verhalten zu dessen Handhabung extrahiert werden können. Anschließend findet ein Abgleich mit den bestehenden Einstellungen und Überzeugungen des Empfangenden statt. Wenn die Person nicht einverstanden ist, muss die Warnung ausreichend überzeugend sein, um eine Einstellungsänderung in Richtung Zustimmung hervorzurufen. Schließlich sollte die Warnung zu einem korrekten Einverständnisverhalten motivieren. Feedbackschleifen und Rückkopplungen von späteren Stufen auf vorherige Stufen zeigen, dass spätere Schritte vorherige positiv als auch negativ beeinflussen können. So haben Einstellungen und Überzeugungen beispielsweise einen Effekt auf die Empfänglichkeit und die Aufmerksamkeit für bestimmte Themen.

Aus dem Modell haben Conzola und Wogalter (2001) Attribute für gut gestaltete Warnhinweise abgeleitet, die das erfolgreiche Überwinden der Stufen hin zur nächsten begünstigen. Quellen oder Sendende sollen positiv assoziiert und vertrauenswürdig sein, sodass der Information Glauben geschenkt wird. Die Warnung soll sich salient vom Hintergrund abheben, was einen Wechsel der Aufmerksamkeit auf Warnung begünstigt, sowie sinnvoll im

Sichtfeld der Zielgruppe und in lokaler Nähe zur möglichen Gefahr platziert sein. Eine ansprechende, ästhetische Gestaltung der Warnung spielt dabei eine Rolle, die Aufmerksamkeit solange aufrecht zu erhalten, bis die Warnung verstanden wurde. Die Botschaft soll, wenn sie in Schriftform abgebildet ist, so kurz wie möglich und so lang wie nötig sein. Hier muss ein sinnvoller Kompromiss zwischen Kürze und Vollständigkeit gefunden werden. Zur Förderung des Verständnisses und der eigenen Betroffenheit muss ein real greifbares Risiko explizite vermittelt werden. Ausgehend vom niedrigsten Fähigkeitsniveau muss der Inhalt für alle Personen der Zielgruppe verständlich sein. Bildsymbole können sowohl zur Vervollständigung von Warntexten verwendet werden als auch alleine abgebildet werden, wenn ein allein stehendes Bildsymbol leicht erkannt und verstanden wird. Dies fördert zum einen die multilinguale und internationale Verständlichkeit, erweitert die Zielgruppe aber auch um Personen, die nicht lesen können wie Analphabeten oder Kinder im Vorschulalter. Schriftzüge sind im Umkehrschluss dann sinnvoll, wenn die Zielgruppe definitiv aus einer Sprachregion kommt und lesen kann. Testphasen sind im Zusammenhang mit der Nutzung von Bildsymbolen in jeden Fall sinnvoll, um die Eindeutigkeit des Zeichens zu validieren. Die Botschaft soll grundsätzlich gängige Einstellungen und Überzeugungen der Zielgruppe ansprechen und derart überzeugend sein, dass konträre Einstellungen und Überzeugungen in Richtung Zustimmung überdacht werden. Letztlich spielen bei der Verhaltensumsetzung noch Kosten-Nutzen-Erwägungen des Empfangenden eine Rolle: Lohnt sich der Aufwand für das Verhalten entsprechend des Benefits? Hierbei werden motivationale Aspekte angesprochen, die entsprechende der Fehlerklassifikation von Reason (1990) zu den bewusst ausgeübten Verstößen führen können. An dieser Stelle hilft die Protection Motivation Theory (PMT) in der erweiterten Version von Rogers, Cacioppo und Petty (1983) dabei, die Rolle der Motivation bei der Bildung einer Handlungsintention nach einer Warnung zu verstehen, die anschließend vorgestellt wird.

Insgesamt helfen die Überlegungen von Conzola und Wogalter (2001) zu den Attributen gut gestalteter Warnungen dabei, die getesteten Maßnahmen post hoc zu evaluieren und Argumente für die Bildung gerichteter Hypothesen zu entwickeln. Daher werden sie im späteren Kapitel 2.4 explizit wieder aufgegriffen.

2.2.4. Protection Motivation Theory (PMT) von Rogers, Cacioppo und Petty (1983).

Die PMT dient als Modell dafür, wie sich Furchtappelle auf die Schutzmotivation i.S. der Intention auswirken, sich gesundheitsförderlich zu verhalten. Ursprünglich wurde die Theorie entwickelt, um die Wirkung von beispielweise Nichtraucherkampagnen oder die Wahrnehmung von privaten Vorsorgeuntersuchungen im Gesundheitssektor zu verbessern. Die PMT

ist in Abbildung 4 dargestellt und wird im Folgenden anhand von Beispielen bei der Querung eines BÜ erläutert, da sie zur Erklärung von bewusst begangenen Verstößen genutzt werden kann. Eine absichtlich unsichere Querung kann Folgen für die eigene Gesundheit und sogar das eigene Leben haben, weshalb die im Modell postulierte Schutzmotivation auch am BÜ eine Rolle spielt.

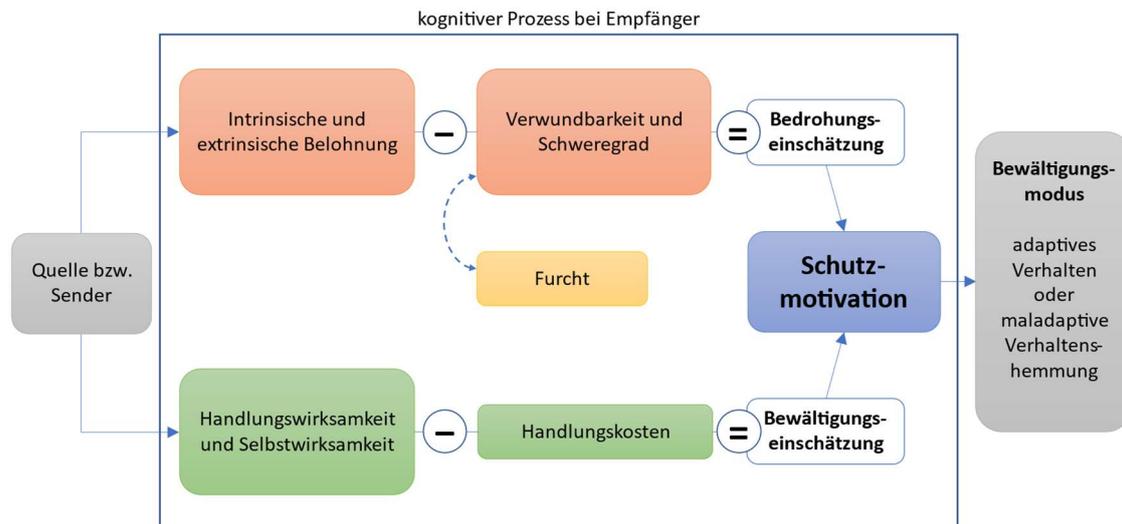


Abbildung 4 Protection Motivation Theory (in Anlehnung an Rogers et al., 1983, S. 168, Lippke und Renneberg, 2011, S. 39).

Entsprechend der Theorie von Rogers et al. (1983) geht von einer Quelle oder einem Sendenden eine Furchtappell-Botschaft an den Empfangenden aus (Ein Fahrradfahrender nimmt die Warnung vor dem nahenden BÜ visuell wahr). Diese Botschaft enthält vier wesentliche Komponenten, die im Rahmen der kognitiven Informationsverarbeitung bewertet werden:

1. die Schwere der Bedrohung,
2. die eigene Betroffenheit des Empfangenden,
3. die Effektivität der vorgeschlagenen Bewältigungsstrategie,
4. die Selbstwirksamkeit des Empfangenden i.S. der individuellen Einschätzung darüber, ob das gewünschte Verhalten umsetzbar ist.

Aus diesen vier Komponente können die relevanten Konstrukte des kognitiven Prozesses für die Wahl des Bewältigungsmodus abgeleitet werden. Die erste wesentliche Modellkomponente ist die Bedrohungseinschätzung (Wie gefährlich ist der nahende BÜ für mich?). Sie wird als eine Funktion aus der erwarteten extrinsischen und intrinsischen Belohnung des bisherigen ungewünschten Verhaltens (extrinsisch: Wenn ich zu spät nach Hause komme, ist meine Frau böse, daher rase ich ohne die Geschwindigkeit zu reduzieren und nach einem Zug

zu schauen; intrinsisch: Wer bremst verliert! Ich bin kein Verlierer.) sowie der Verwundbarkeit i.S. der Eintrittswahrscheinlichkeit des Ereignisses (Abwägung, wie oft wohl ein Zug kreuzt) und dem Schweregrad der Folgen (Eine Kollision mit einem Zug werde ich nicht überleben) beschrieben. Rogers et al. (1983) integrieren noch eine entscheidende emotionale Komponente in ihr Modell: die Furcht. Empfindet der Empfangende die potenzielle Bedrohung als gravierend, entsteht Furcht, was in einer reziproken Reaktion dazu führt, dass die Bedrohung als noch schwerwiegender eingeschätzt wird. Parallel zur Bedrohungseinschätzung findet die Bewältigungseinschätzung statt, welche die zweite relevante Modellkomponente darstellt. Sie ergibt sich aus der Funktion der Handlungswirksamkeit (Wenn ich die Geschwindigkeit reduziere, kann ich auch ausgiebig nach einem Zug gucken, wodurch ich einen Unfall sicher vermeide) und der Selbstwirksamkeit (Ich bin grundsätzlich dazu in der Lage, zu bremsen und ausgiebig zu schauen) sowie der Handlungskostenkomponente (Bremsen in dieser Situation ist für mich ungewohnt und unbequem, da ich bremsen generell vermeide. Wer bremst, verliert nicht nur das Rennen, sondern auch Energie!). Die Bedrohungseinschätzung und die Bewältigungseinschätzung zusammen ergeben die Intention im Sinne der beschriebenen Schutzmotivation. Sie führt dazu, dass entweder ein gewünschtes, adaptives Verhalten gezeigt wird oder aber, dass eine maladaptive Hemmung des sicheren Verhaltens stattfindet. Besonders positiv hervorzuheben bei der PMT ist, „dass sie neben [der] Bedrohungseinschätzung verschiedene personale Ressourcen (Bewältigungseinschätzung) berücksichtigt“ (Lippke & Renneberg, 2011, S. 40).

Ein Aspekt, der sich aus der PMT für die Entwicklung effektiver Maßnahmen zur Förderung von sicherem Querungsverhalten am BÜ ergibt, ist, dass die bereits von Conzola und Wogalter (2001) beschriebene bedeutsame Komponente der Motivation durch die Maßnahme gefördert werden sollte. Hierbei kann vor allem die Wahrscheinlichkeit für bewusst begangene Verstöße verringert werden. Die extrinsischen und intrinsischen Vorteile durch einen Verstoß sollten durch die Verdeutlichung des eigenen Risikos abgeschwächt werden. Ebenso sollten die Handlungs- und Selbstwirksamkeit durch die Betonung der Einfachheit der Bewältigungsstrategie gestärkt werden. Daraus ergibt sich eine hohe Schutzmotivation i.S. der Intention, nach einem Schienenfahrzeug Ausschau zu halten und dafür auch die eigene Geschwindigkeit bei der Annäherung an den BÜ zu verringern. Insgesamt ist davon auszugehen, dass alle Straßenverkehrsteilnehmenden grundsätzlich eine hohe Schutzmotivation dahingehend aufweisen, den BÜ sicher zu queren. In einer Studie untersuchten Glendon und Walker (2013) die Wirkung von Anti-Speeding-Botschaften, die entsprechend der Kernkom-

ponenten der PMT formuliert wurden. Für die Komponente *Schweregrad* wurden beispielsweise die Aussagen *Kill your speed, not yourself*, und *Don't speed. Don't kill your mates* oder für die Komponente *Selbstwirksamkeit* die Aussagen *You can stay safe. You can drive the speed limit* und *You can save a life. You can drive the speed limit* auf einer fünfstufigen Skala dahingehend bewerten, ob sie in Folge der Wahrnehmung einer solchen Botschaft während einer Autofahrt ihre Geschwindigkeit kontrollieren und ggf. anpassen würden. Die Ergebnisse bestätigen, dass Botschaften mit von PMT abgeleiteten Inhalten zu durchweg höheren berichteten Absichten führten, innerhalb des Geschwindigkeitslimits zu fahren als allgemeinere Warnhinweise wie *Every K over is a killer* oder *Drive within the speed limit*.

Grundsätzlich kann aus der PMT und daraus resultierender Studienergebnisse geschlossen werden, dass die Maßnahmen zur Vermeidung von bewussten Verstößen bei der Querung am BÜ besonders effektiv im Sinne einer Warnbotschaft sind, wenn sie entsprechend der Komponenten der PMT formuliert sind. Dies würde für Verhalten gelten, das absichtlich geschieht und i.d.R. den geringsten Anteil von unsicherem Verhalten am BÜ ausmacht wie beispielsweise das Stehenbleiben auf dem BÜ, um ein Foto zu machen oder das schnelle Passieren, obwohl bereits ein Schienenfahrzeug in Sichtweite wahrgenommen wurde. Dennoch ist der Aspekt der Schutzmotivation und der Intention, sich regelkonform zu verhalten im Straßenverkehr allgegenwärtig, um sich selbst und andere nicht unnötig zu gefährden.

2.2.5. Anknüpfungspunkte für Interventionen.

Zusammengefasst können auf Basis der beschriebenen Theorien, Modelle und Studien drei Hauptursachen für Fehlverhalten am nicht-technisch gesicherten BÜ benannt werden, an denen Interventionen zur Förderung von sicherem Verhalten anknüpfen sollten:

1. Die kognitive Informationsverarbeitung ist nicht tief genug, sodass die BÜ-Infrastruktur als solche zwar erkannt wird, aber durch Defizite in der Aufmerksamkeit oder beim Gedächtnis entsprechende Schemata und Verhaltensskripte nicht abgerufen werden können. Hierbei könnten kurze Hinweise wie beispielsweise Prompts helfen, tiefere Verarbeitungsschritte zu aktivieren und die Straßenverkehrsteilnehmenden zu adäquatem Verhalten anzuregen.
2. Bei den Straßenverkehrsteilnehmenden sind entsprechende Schemata und Verhaltensskripte nicht vorhanden, da sie beispielsweise mit nicht-technisch gesicherten BÜ nicht vertraut sind oder sich ein „Kein-Zug-Schema“ etabliert hat, da noch nie ein Schienenfahrzeug in dem Setting wahrgenommen wurde. Das Fehlverhaltens

entsteht dementsprechend auf Grund von Unwissenheit oder durch eine maladaptive Reaktion. An dieser Stelle helfen konkrete Warnhinweise oder Informationen, die zur bewussten Auseinandersetzung mit der Situation anregen, um im Rahmen einer adaptiven Problemlösung die Herausforderung des Querens erfolgreich zu bewältigen.

3. Die Motivation der Straßenverkehrsteilnehmenden ist zu gering, sich im Straßenverkehr grundsätzlich angemessen zu verhalten. Dadurch werden bewusst Verstöße in Kauf genommen. Warnappelle, die die Bedrohungseinschätzung erhöhen, helfen an dieser Stelle, um die eigene Betroffenheit durch Fehlverhalten zu verdeutlichen und die Schutzmotivation zu erhöhen.

2.3. Getestete Maßnahmen

Grundsätzlich empfehlen Matthies, Kastner, Klesse und Wagner (2011) im Rahmen umweltsychologischer Interventionen ein strategisches Vorgehen bei der Planung und Entwicklung. Unverzichtbar sind hierbei die Analyse des Ist-Standes, eine präzise Definition der Interventionsobjekte und der Zielgruppe sowie eine theoriegeleitete Anpassung der Intervention an die Eigenschaften der Zielgruppe, was im Rahmen des Projektes SAFER-LC geschehen ist. Nach Dressler und Silla, et al. (2020) wurde zunächst ein Pool von Gestaltungsideen auf Basis der Forschungsliteratur sowie eines Expert:innenworkshops angelegt. In einem zweiten Schritt wurden die gesammelten Maßnahmen kategorisiert und eingestuft, sowie hinsichtlich ihrer Aussichten zur Verringerung des Unfallrisikos bewertet. Der daraus resultierende Satz umfasst 89 Maßnahmen und steht als webbasierte Toolbox für Praktiker:innen im Bereich der Straßen- und Eisenbahnsicherheit zur Verfügung. Zwei vielversprechende, innovative Maßnahmen wurden an einem nicht-technisch gesicherten BÜ in Braunschweig getestet. Um die in Kapitel 2.4 beschriebenen Argumenten zur Wirksamkeit der getesteten Maßnahmen und zur Hypothesenbildung besser nachvollziehen zu können, werden sie im Folgenden beschrieben.

2.3.1. Beschreibung der Maßnahmen.

Als eine innovative Maßnahme zur Förderung von sicherem Verhalten am BÜ bei der Querung wurde die Straßenmarkierung *Kommt ein Zug?* mit Pfeilen links und rechts bewertet (siehe Abbildung 5).



Abbildung 5 Erste getestete Maßnahme: Straßenmarkierung (Silla et al., 2019, S. 126).

Die Markierung wurde 35 m vor dem nicht-technischen BÜ mit reversibler Kreidefarbe aufgetragen. Die Frage danach, ob sich ein Zug nähert, spricht die Querenden direkt an und regt zur Reflektion der Umgebung an. Personen, die unwissend angesichts des nahenden BÜ sind, werden auf die potenzielle Gefahrenquelle aufmerksam gemacht. Ihnen wird der Impuls dazu gegeben, eine adäquate Handlungsstrategie für die erfolgreiche, sichere Überquerung zu entwickeln. Ortskundige Personen werden hingegen durch den kurzen Fragesatz an richtiges Querungsverhalten erinnert. Im Rahmen gängiger Handlungsmodelle wie z.B. der *Theory of Planned Behaviour* (Ajzen, 1991) spielt die Perceived Behavior Control (PBC) eine essenzielle Rolle, da sie sowohl direkt oder auch indirekten Einfluss auf das Verhalten ausübt: Die Ausführung eines Verhaltens wird wahrscheinlicher, wenn die Person davon überzeugt ist, die Umsetzung eines bestimmte Verhaltens zu bewältigen und denkt, über die „requisite opportunities and resources“ (Ajzen, 1991, S. 182) zu verfügen. Wenn die Querenden durch die Straßenmarkierung direkt danach gefragt werden, ob ein Schienenfahrzeug komme, wird die beschriebene PBC aktiviert. Ihnen wird in Form der Frage deutlich gemacht, dass sie verantwortlich für sich und die Überquerung sind. Da die Beantwortung der Frage für den Großteil der Zielgruppe durch ein niedrigschwelliges Verhalten leicht möglich ist, sollten sowohl die Intention zur Verhaltensausführung als auch das Verhalten an sich gefördert werden.

Die zweite getestete Maßnahme ist das in Abbildung 6 dargestellte gelbes Blinklicht, das bei Annäherung aktiviert wird.

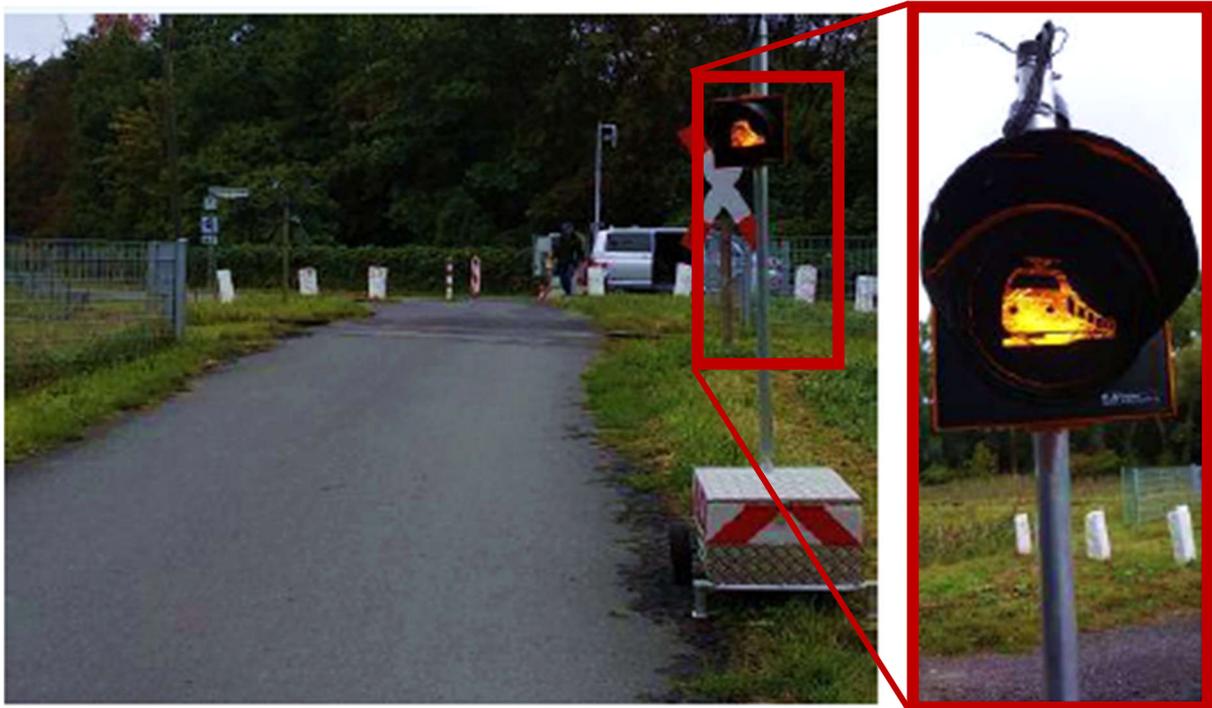


Abbildung 6 Zweite getestete Maßnahme: Blinklicht (nach Silla et al., 2019, S. 127). Links: Blick aus nordöstlicher Richtung auf den BÜ. Rechts: Großaufnahme des Blinklichts.

Als Icon wurde das aus diversen Verkehrszeichen gängige Bild eines herannahenden Zuges gewählt (z.B. Zeichen 151 „Bahnübergang“, Anlage 1 zu § 40 Absatz 6 und 7 der StVO, Bundesministerium der Justiz und für Verbraucherschutz, (2020). Cian, Krishna und Elder (2015) nutzten für eine Studie zur Gestaltung von Warnzeichen, dass Menschen mentale Bilder erzeugen können: Durch die Gestaltung von dynamischen Objekten auf den Schildern, die den Eindruck entstehen lassen, dass sich im Bild eine Bewegung abspielt, zeigten die Versuchspersonen eine frühere Aufmerksamkeit auf die Gefahrensituation, eine höhere Wachsamkeit in der Gefahrensituation und eine schnellere Reaktion auf Anzeichen von Gefahr. Das Zeichen des Zuges bei dem Blinklicht impliziert Bewegung, da der Zug sich nach hinten verbiegt und dadurch auf den Betrachtenden zu zukommen scheint, was der Studie entsprechend die Wirkung der Maßnahme auf das gewünschte Verhalten unterstützen sollte. Hervorzuheben ist die Salienz des Blinklichtes, da es sich deutlich vom Hintergrund abhebt und durch die Aktivierung bei Annäherung eine interaktive Komponente enthält. Kritisch anzumerken ist jedoch, dass unabhängig voneinander Personen Zweifel an der Verständlichkeit und Eindeutigkeit der Botschaft geäußert haben: Das Blinklicht könnte als konkrete Warnung vor einem nahenden Zug verstanden werden ähnlich wie es bei technisch-gesicherten BÜ der Fall ist, bei denen eine rot blinkende Signalanlage über dem Andreaskreuz ein kreuzendes Schienenfahrzeug ankündigt.

2.3.2. Interventionspsychologische Einordnung.

Beide getestete Maßnahmen gehören nach Klöckner (2015) zu den kommunikationsbasierten Strategie und können in die Gruppe der *Prompts* eingeordnet werden. Prompts zeichnen sich durch eine kurze, geschriebene Nachricht oder ein Zeichen aus, die in einer spezifischen Situation ein bestimmtes Verhalten hervorrufen sollen (Abrahamse & Matthies). Der Gebrauch von Prompts bietet sich an, wenn die Zielgruppe dem Zielverhalten gegenüber grundsätzlich positiv eingestellt ist, „but lacks a cue in the situation where the behaviour is required. Thus, prompts can be assumed to overrule the automatic elicitation of a problematic behaviour“ (Steg, 2013, S. 311). Klöckner (2015) fasst zusammen, dass Prompts dann effektiv sind, wenn sie in unmittelbarer Nähe des Ortes des Zielverhaltens platziert werden, wenn sie möglichst konkrete Anweisungen darüber enthalten, was das Zielverhalten ist und wenn sie gut sichtbar sind. Die Wirkung von Prompts lässt jedoch nach, wenn sich die Zielgruppe an sie gewöhnt, was durch dynamische Prompts vermieden werden könnte. In einer verkehrspsychologischen Studie von Cox, Cox und Cox (2000) wurde die Anschnallrate von Autofahrenden an der Ausfahrt eines Parkplatzes mit einem Stoppschild untersucht. Nach der Kontroll-erhebung, die zu zwei Zeitpunkten stattfand, wurde ein Prompt mit der Aufschrift *Buckle up, stay safe* direkt unterhalb des Stoppschildes angebracht, um die Autofahrenden an das Anschnallen zu erinnern. Cox et al. (2000) erhoben eine Woche nach der Installation die Anzahl der Personen, die angeschnallt von dem Parkplatz fuhren. Das Ergebnis war ein signifikanter Anstieg der Anzahl der Personen, die angeschnallt den Parkplatz verließen, von 68 % und 72 % zum Zeitpunkt der beiden Kontrollerhebungen auf 94 % zur Testbedingung mit dem Schild. Darüber hinaus schnallten sich durch den Prompt signifikant mehr Personen am Stoppschild an als zur Kontrollerhebung, was die Wirksamkeit der Botschaft als kurze Erinnerung unterstützt. Ebenso wie der Prompt an der Ausfahrt bei Fahrtantritt sollten sowohl die Straßenmarkierung als auch das Blinklicht als Prompt den Straßenverkehrsteilnehmenden zu adäquatem Verhalten erinnern und anregen.

Neben der Funktion der Maßnahmen, dass sie bei den Personen im Sinne eines Prompts ein spezifisches Verhalten hervorrufen sollen, sind die getesteten Maßnahmen eine Warnung vor einer potenziellen Gefahr. Somit haben die beiden Maßnahmen auch eine informierende Funktion ähnlich wie bei der Interventionsgruppe der *Informationsbereitstellung* nach Klöckner (2015), wobei sie nach dessen Definition nicht den Detaillierungsgrad einer aufklärenden Kampagne wie bei umweltpsychologischen Interventionen haben. Übertragbar ist jedoch das eigentliche Ziel von aufklärenden Kampagnen auf die Warnbotschaft: „The provision of information is built on the assumption that knowledge ... make people change

their behaviour. ... Providing information might furthermore be understood as a technique of attitude change“ (Klößner, 2015, S. 164). Conzola und Wogalter (2001) beschreiben im bereits vorgestellten C-HIP-Modell neben der Wirkung von Warnungen deren informierende Funktion: Demnach werden Warnhinweise zur Vermittlung bedeutsamer, sicherheitsrelevanter Botschaften angebracht, um ein sicheres Verhalten zu fördern sowie unsicheres Verhalten zu reduzieren. Die Warnung kann nach deren Auffassung derart überzeugend sein, dass vorhandene Einstellungen überdacht werden.

2.4. Bewertung der Maßnahmen anhand der theoretischen Herleitung

Anhand der in 2.1 und 2.2 beschriebenen Theorien zum Verhalten im Straßenverkehr als regelbasierte Reiz-Reaktions-Verknüpfung und zur Klassifikation sowie zu möglichen Ursachen von Fehlverhalten werden die getesteten Maßnahmen im Rahmen der vorliegenden Masterarbeit bewertet. Das Hauptaugenmerk liegt auf den von Conzola und Wogalter (2001) abgeleiteten Attributen aus dem C-HIP-Modell einer guten Warnung sowie aus der Einschätzung der Maßnahmen auf Basis der PMT von Rogers et al. (1983), die im Wesentlichen die genannten hypothetischen Wirksamkeitsfaktoren anderer genannter Autor:innen abdecken. Ziel ist hierbei die Generierung von gerichteten Hypothesen für die Datenauswertung.

Für jedes Attribut einer gut gestalteten Intervention wird geschaut, ob es in der Maßnahme als erfüllt oder nicht erfüllt beurteilt werden kann, oder ob eine eindeutige Beurteilung nicht stattfinden kann. Das Ergebnis ist in Abbildung 7 dargestellt und wird im Folgenden näher erläutert.

Theoretischer Hintergrund	Attribut		
C-HIP-Modell (Conzola & Wogalter, 2001)	vertrauenswürdige Quelle	+	+
	hohe Salienz	-	o
	Platzierung im Raum entspricht Sichtfeld der Zielgruppe	o	+
	gelungener Kompromiss zwischen Kürze und Vollständigkeit der Botschaft	+	o
	gegebene lokale Nähe zur Gefahr	+	+
	Ästhetik des Maßnahmendesigns	+	+
	möglichst explizite Vermittlung eines real greifbaren Risikos	o	o
	für alle Personen der Zielgruppe verständlich	-	o
	sinnvolle Auswahl/ Ergänzung von Bildsymbolen und Schrift	-	o
	spricht grundsätzlich gängige Einstellungen und Überzeugungen an	+	+
	Bilanz der Kosten/Nutzen -Abwägungen spricht für Verhaltensausführung	+	+
	PMT (Rogers et al., 1983)	erhöht Bewältigungseinschätzung	o
erhöht Bedrohungseinschätzung		-	o

Abbildung 7 Beurteilung der Maßnahmen anhand der von Conzola und Wogalter (2001) beschriebenen Attribute für die Gestaltung effektiver Warnhinweise sowie anhand der im PMT von Rogers et al. (1983) beschriebenen Kernkomponenten. Rotes Minuszeichen: nicht erfüllt; grünes Pluszeichen: erfüllt; schwarzer Kreis: nicht eindeutig beurteilbar.

Beide Maßnahmen stammen offensichtlich eher aus vertrauenswürdigen Quellen, da die Veränderung an der Infrastruktur eines BÜ ferner verboten wäre. Die lokale Nähe zur potenziellen Gefahr ist ebenfalls gegeben. Grundsätzlich sind die Maßnahmen ästhetisch gestaltet, sodass die Aufmerksamkeit der Personen solange aufrecht erhalten bleiben sollte, dass alle Informationen zum Verständnis extrahiert werden können. Die Botschaft beider Maßnahmen spricht die grundsätzlich gängige Einstellung und Überzeugung der Zielgruppe an, den BÜ sicher zu überwinden. Hierbei sollte die Bilanz der Kosten/ Nutzen-Abwägungen generell für die Verhaltensausführung sprechen, da sicheres Verhalten am BÜ wenig Aufwand erfordert und angemessen angesichts der eigenen Gesundheit sein müsste. Die Explizitheit der beiden Maßnahmen bezüglich des möglichen Risikos ist als mittelmäßig und steigerungsfähig zu beurteilen, da beispielsweise auch die fatalen Folgen von Unfällen bildlich gezeigt werden könnten.

Unterschiede zwischen den Maßnahmen ergeben sich bezüglich der Salienz. Wolfe und Horowitz ordneten 2004 in einer Zusammenfassung bisheriger Studienergebnisse die Eigenschaften von Stimuli entsprechend ihrer Fähigkeit, die visuelle Aufmerksamkeit auf sich zu ziehen. Das Ergebnis war, dass insbesondere die Farbe, Bewegung und Größe von visuellen Stimuli „undoubted attributes“ (Wolfe & Horowitz, 2004, S. 6) sind. Bei diesen drei Eigenschaften sind Vorteile seitens des Blinklichtes zu erwarten, da es ein farbiges Leuchtsignal

ist, das größer im Sichtfeld platziert ist. Der Schriftzug ist auch nach Einschätzung der Studienbegleitenden dem Blinklicht bei diesem Attribut unterlegen. Durch starke Verschmutzungen der Fahrbahn (beispielsweise durch landwirtschaftlichen Verkehr) oder Laubfall der Bäume ist die Lesbarkeit der Straßenmarkierung gefährdet, vor allem dann, wenn keine regelmäßige Straßenreinigung auf abgelegeneren Strecken stattfindet. Darüber hinaus befindet sich die Straßenmarkierung außerhalb des direkten Sichtfelds der Querenden, die beispielsweise gerade auf den Gegenverkehr achten. Weitere Nachteile ergeben sich bei der Verständlichkeit der Markierung, da Analphabeten, Personen, die kein Deutsch verstehen, Kinder im Vorschulalter etc. die Schrift nicht verstehen. Auch das Blinklicht ist durch die beschriebene Fehlinterpretation nicht eindeutig verständlich und das Zugsymbol allein irreführend: Die Personen könnten denken, das Blinklicht fungiere als konkrete Warnung vor einem nahenden Schienenfahrzeug. Durch die direkte Ansprache der Personen bei der Markierung in Form einer leicht zu beantwortenden Frage wird die Bewältigungseinschätzung der Personen derart erhöht, dass die nahende potenziell gefährliche Situation als leicht kontrollierbar empfunden wird. Bei der Beurteilung der Bedrohungseinschätzung ergeben sich leichte Vorteile bei dem Blinklicht gegenüber der Markierung: Das Wort *Zug* macht in dem Fragesatz zwar deutlich, was konkret die Gefährdung ist, jedoch wird das Wort weder besonders hervor gehoben noch durch Farben oder Bilder die Schwere der Bedrohung betont. Im Gegensatz dazu ist das warnend gelb blinkende Zugsymbol förderlicher. Das Blinklicht bietet jedoch keinerlei Anreize für die erfolgreiche Bewältigung der nahenden Situation, wobei jedoch von einem gesunden Menschenverstand erwartbar ist, dass mindestens ein Kontrollblick nach einem Zug in Folge eines solchen Warnappells ausgehen müsste.

Insgesamt weisen beide Maßnahmen Vor- und Nachteile auf. Das Fazit aus den Überlegungen aus Abbildung 7 zeigt einen Vorteil für das Blinklicht (sechs Pluszeichen, ein Minuszeichen, sechs neutrale Bewertungen) gegenüber der Markierung (sechs Pluszeichen, vier Minuszeichen, drei neutrale Bewertungen). Daher wird bei der Hypothesenformulierung in der vorliegenden Masterarbeit davon ausgegangen, dass die Wahrscheinlichkeit für sicheres Verhalten am BÜ durch die Maßnahme des Blinklichts höher ist als durch die Maßnahme der Straßenmarkierung.

2.5. Gewöhnungseffekt

Die Wirksamkeit der getesteten Maßnahmen könnte sich über die Zeit verändern. Sowohl im umweltsychologischen Bereich bei Prompts und Nudges als auch in der Forschung zur Wirksamkeit von Warnbotschaften wird davon ausgegangen, dass die Effektivität dieser Art von

Maßnahmen über die Zeit abnimmt (Abrahamse & Matthies; Conzola & Wogalter, 2001; Steg, 2013; Wogalter & Laughery, 1996). Die Annahme dahinter ist, „that when a warning becomes habituated over time from repeated exposures, attention is less likely to be allocated to the warning on subsequent occasions“ (Conzola & Wogalter, 2001, S. 313). Die Habituation ist nach den Neurowissenschaftlern Groves und Thompson (1970) definiert als verminderte Reaktion auf wiederholte Stimulation. Zwei wesentliche Effekte, die im Zusammenhang mit der Habituation erforscht wurden, sind für den zeitlichen Verlauf der Entwicklung eines neuronalen Gewöhnungseffektes an einen Stimulus in einer spezifischen Situation maßgeblich: “the frequency effect (more rapid and pronounced habituation with more rapid frequency of stimulation) and the intensity effect (the weaker the stimulus the more rapid and pronounced the habituation)” (Thompson, 2009, S. 132). Wann die Gewöhnung an einen Prompt oder eine Warnung konkret einsetzt, kann somit nicht genau vorhergesagt werden und ist theoretisch ab der ersten Präsentation möglich, wenn der Stimulus reizarm ist.

Eine längerfristige Erprobung der Maßnahmen konnte im Projekt SAFER-LC nicht umgesetzt werden, so dass der Testaufbau die Untersuchung möglicher Gewöhnungseffekte, wie sie in einem dauerhaften Einsatz eintreten könnten, nicht erlaubt. Da sich die Testzeiträume jedoch zumindest über zwei Wochen erstreckten und wenigstens bei einem Teil der den Bahnübergang nutzenden Personen eine wiederholte Nutzung zu erwarten ist, soll dennoch geprüft werden, ob bereits in diesem kurzen Einsatzzeitraum Hinweise auf mögliche Gewöhnungseffekte zu beobachten sind.

Auf Grund der zentrumsnahen Lage des Bahnübergangs in der Großstadt Braunschweig und den gewählten Tageszeiten an den Messtagen (morgens und nachmittags) ist davon auszugehen, dass viele zu Fuß Gehende und Fahrradfahrende auf ihren regelmäßigen Wegen (bspw. zur Arbeit) beobachtet wurden. Zur Sondierung möglicher Effekte der Dauer des Maßnahmeneinsatzes werden in der vorliegenden Masterarbeit jeweils ein Tag zu Anfang und ein Tag zu Ende jeder Testphase miteinander verglichen. Da es sich hierbei um unterschiedliche Wochentage handelt (Dienstag vs. Donnerstag), wird ebenfalls überprüft, ob ein Wochentageffekt vorliegt, der gegebenenfalls eine nicht gefundene Veränderung über die Zeit verschleiert oder eine gefundene erzeugt haben könnte.

Wie oben erwähnt, hat die Überprüfung eines möglichen Gewöhnungseffekts lediglich explorativen Charakter. Würde sich die Wirksamkeit der Maßnahmen bereits über die kurzen Testzeiträume hinweg verringern, wäre dies ein deutlicher Hinweis auf Optimierungs-

bedarf. Hingegen kann aus einem Ausbleiben eines solchen Effekts in der vorliegenden Untersuchung nicht geschlossen werden, dass keine Gewöhnung erfolgt. Hierfür ausschlaggebend sind folgende Sachverhalte:

1. Die Häufigkeit, mit der die beobachteten Personen den Bahnübergang zum Zeitpunkt der Beobachtung bereits zuvor überquert hatten, sowie die zeitliche Abfolge möglicher vorheriger Querungen sind unbekannt. Zwar ist es wahrscheinlich, dass ein Teil der zum späten Messzeitpunkt beobachteten Personen den beiden Maßnahmen bereits mehrfach zuvor begegnet waren. Es kann jedoch anhand der vorliegenden Daten die Möglichkeit nicht ausgeschlossen werden, dass alle Beobachteten den Bahnübergang mit den Maßnahmen an diesem Tag zum ersten Mal überquerten.
2. Die Testdauer von zwei Wochen je Intervention ist verhältnismäßig kurz.

Die Beobachtung eines Gewöhnungseffektes ist durch die Studienlage und die Eigenschaften der getesteten Maßnahmen dennoch grundsätzlich möglich, so dass bei der Fragestellung und zugehörigen Hypothese von dessen Existenz ausgegangen wird.

2.6. Auswahl und Objektivität der Datenerhebungsmethode

Feldbeobachtungen gehören nach Eby (2011) in der Verkehrspsychologie zu den Standardmethoden, da hierbei direkt das Verhalten in der natürlich auftretenden Situation ohne Verzerrungen erfasst wird. Hierbei ist zu beachten, dass Beobachtungsverfahren lediglich den Zugang zur Außensicht von Handlungen und Ereignissen ermöglichen ohne die genauen Bedeutungen, Intentionen und Ziele zu erfragen (Hussy et al., 2010). Nach Eby (2011) sind Feldbeobachtungen jedoch vor allem dann vorteilhaft, wenn vier Aspekte zutreffen:

1. Die Häufigkeit des Auftretens eines Ereignisses sind von Interesse.
2. Das Verhalten ist akkurat und reliabel von außen beurteilbar.
3. Die beobachteten Daten müssen repräsentativ sein und die Stichprobengröße entsprechend groß unter der Beachtung von ökonomischen Aspekten bei Erhebung und Auswertung.
4. Es stehen ausreichend Ressourcen zur Verfügung.

Demuth fasst zusammen, dass „viele Aspekte menschlichen Handelns [...] subtil, flüchtig und leicht zu übersehen“ sind, wenn sie in situ annotiert werden. Die Videotechnologie ermöglicht die über einen Zeitraum aufgenommenen Daten detailliert auszuwerten, bei Unklarheiten zurück zu spulen und die Daten von mehreren Beobachtenden aus der gleichen

Blickperspektive auswerten zu lassen. Aus den genannten Vorteilen einer nicht-teilnehmenden, vermittelten, systematischen Feldbeobachtung entschied sich das Projektteam des DLR für den Einsatz eines Verkehrsüberwachungssystems, mit dem gering aufgelöste und somit anonymisierte Videoaufnahmen realisiert werden konnten.

Das generelle Ziel der experimentellen Untersuchungsplanung sollte sein, eine Operationalisierung zu wählen, die eine möglichst objektive Datengewinnung gestattet (Bortz & Schuster, 2010). Hierfür ist entscheidend, dass das Untersuchungsmaterial eindeutig quantifizierbar ist. Dabei spielen sowohl die Qualität des Videomaterials eine entscheidende Rolle, als auch die Genauigkeit der Operationalisierung des Zielverhaltens. Eby (2011) empfiehlt, dass bei der Auswertung von Videomaterial mehrere geschulte Untersuchungsauswertende die Annotation durchführen sollen, um die Generalisierbarkeit der Befunde zu verbessern und Verzerrungen durch die Auswertenden zu vermeiden. Hussy et al. (2010) sehen für das Kriterium der Objektivität bei der Datenerhebung die Interrater-Übereinstimmung als zentrales Konzept. Die Auswertung der Daten des Verkehrsüberwachungssystems fand bereits im Rahmen einer Abschlussarbeit statt und wird in der vorliegenden Masterarbeit repliziert. Die von Bortz und Schuster (2010) empfohlene hinreichende Standardisierung des für die Annotation relevanten Zielverhaltens sollte in der durchgeführten Studie in jedem Fall gegeben sein, da die Variable des Blickverhaltens, operationalisiert als Kopfbewegung nach links und/ oder rechts, eindeutig ist und bei der Videoanalyse wenig Interpretationsspielraum lässt: Findet ein Kontrollblick i.S. der Kopfbewegung nach links und/ oder rechts oder gar nicht statt?

2.7. Hypothesen

Die Anliegen der angestrebten Masterarbeit sind demnach, die Wirksamkeit der entwickelten Maßnahmen zur Förderung des adäquaten Kontrollblickverhaltens an nicht-technisch gesicherten BÜ empirisch zu prüfen und statistische Kennwerte zur Objektivität der Videoannotationemethode (Interrater-Reliabilität auf Basis der Ergebnisse eines anderen Annotierenden) zu ermitteln. Ferner wird außerdem untersucht, ob bereits ein Gewöhnungseffekt an die Maßnahmen aufgetreten ist. Für die Analyse, ob die Maßnahmen das Kontrollblickverhalten gefördert haben, wurde in den Testphasen der beiden Maßnahmen sowie in einer Vergleichsphase ohne zusätzliche Maßnahmen das Verhalten der VRU anhand eines Verkehrsüberwachungssystems erfasst. Das System erfasste gering aufgelöste Videoaufnahmen des BÜ. Die wesentliche Säule der Auswertung für die angestrebte Masterarbeit ist die Annotation des Blickverhaltens in Annäherung an den BÜ (*Schaut der VRU links und rechts nach einem*

Zug?). Operationalisiert wird das Kontrollblickverhalten, indem eruiert wird, ob die beobachteten Personen in der Annäherungszone des BÜ eine Kopfbewegung nach links und/ oder rechts zeigen. Hierbei sind für die Auswertungen entweder die differenzierten drei Ausprägungen (gar kein Kontrollblick, Kontrollblick in eine Richtung und Kontrollblick in beide Richtungen) oder das dichotomisierte Kontrollblickverhalten (adäquat in beide Richtungen oder inadäquat in eine oder in gar keine Richtung) relevant.

Um die Untersuchungsziele zu erreichen, wurden auf Basis der theoretischen Herleitung folgende zentrale Fragestellungen formuliert:

- Wie können die getesteten Maßnahmen mit Blick auf die Förderung sicheren adäquaten Kontrollblickverhaltens bewertet werden?
- Wie hoch ist die Objektivität (Interrater-Reliabilität) der gewählten Untersuchungs- und Auswertungsmethode in der vorliegenden Studie unter den gegebenen Bedingungen und mit dem vorhandenen Equipment in Bezug auf die kodierten Personenvariablen (Geschlecht, Alter, Fortbewegungsmittel) und auf die beiden kodierten Varianten des Kontrollblickverhaltens (adäquates und differenziertes Kontrollblickverhalten)?

Darüber hinaus sollen folgende Fragestellungen ebenfalls betrachtet werden:

- Tritt ein Gewöhnungseffekt an die Maßnahmen wie erwartet auf?
- Wie könnten die Maßnahmen zur Förderung des adäquaten Kontrollblickverhaltens optimiert werden?
- Wie könnte die Untersuchungs- und Auswertungsmethode optimiert werden?

Anhand dieser Fragestellungen und aufgrund der theoretischen Herleitung können folgende Hypothesen abgeleitet werden:

H1: In der Vergleichsphase ohne Maßnahmen ist signifikant weniger Kontrollblickverhalten beobachtbar als in der Testphase mit der Straßenmarkierung.

$$P(\text{adäquater Kontrollblick} \mid \text{Baseline}) < P(\text{adäquater Kontrollblick} \mid \text{Markierung})$$

H2: In der Vergleichsphase ohne Maßnahmen ist signifikant weniger Kontrollblickverhalten beobachtbar als in der Testphase mit dem Blinklicht.

$$P(\text{adäquater Kontrollblick} \mid \text{Baseline}) < P(\text{adäquater Kontrollblick} \mid \text{Blinklicht})$$

H3: In der Testphase mit der Straßenmarkierung ist signifikant weniger Kontrollblickverhalten beobachtbar als in der Testphase mit dem Blinklicht.

$$P(\text{adäquater Kontrollblick} \mid \text{Markierung}) < P(\text{adäquater Kontrollblick} \mid \text{Blinklicht})$$

H4: Bei beiden getesteten Maßnahmen tritt ein Gewöhnungseffekt auf.

Je Testbedingung mit Maßnahme:

$P(\text{adäquater Kontrollblick} \mid 1. \text{ Messtag}) > P(\text{adäquater Kontrollblick} \mid 4. \text{ Messtag})$

Kontrolle möglicher überlagernder Wochentageeffekte:

$P(\text{adäquater Kontrollblick} \mid \text{Dienstage}) = P(\text{adäquater Kontrollblick} \mid \text{Donnerstage})$

H5: Die Verhaltensbeobachtungen der beiden Annotierenden sollten nach Landis und Koch (1977) mindestens *mittelmäßig* (engl. *moderate*, $K > .41$) miteinander übereinstimmen.

Cohens K bzw. Yules Y > .41

Die unabhängigen Variablen sind jeweils die Testbedingungen (nominalskaliert): Vergleichsphase ohne Maßnahme, Testphase mit Schriftzug, Testphase mit Blinklicht. Die abhängige Variable sind das Auftreten von Kontrollblickverhalten (nominalskaliert), das auf zwei Arten kodiert wird: (a) adäquates Kontrollblickverhalten - dichotome Merkmalsausprägung: VRU schaut adäquat in beide Richtungen oder schaut inadäquat entweder nur in eine Richtung oder gar nicht. (b) differenziertes Kontrollblickverhalten - drei Kategorien: VRU guckt gar nicht, guckt in eine Richtung oder guckt in beide Richtungen nach einem Schienenfahrzeug.

3. Methoden

Im folgenden Methodenteil wird zunächst das Setting der Feldbeobachtungsstudie beschrieben. Anschließend wird auf die Strichprobengewinnung sowie deren Zusammensetzung eingegangen und der Ablauf der beiden Annotationen erklärt. Es folgt eine nähere Betrachtung der Datenerhebungsmethode, bevor mit der Vorstellung der statistischen Auswertung das Kapitel abgeschlossen wird.

3.1. Beschreibung des Settings

Die Studie fand von Mitte August bis Ende September 2019 in der Großstadt Braunschweig im Randgebiet der Innenstadt statt. Der BÜ ist für vierrädrige Fahrzeuge gesperrt und wird hauptsächlich von zu Fuß Gehenden und Fahrradfahrenden genutzt. Er verbindet zwei angrenzende Stadtteile miteinander und ist daher hoch frequentiert. Die Straße führt durch eine Wiesen- und Waldlandschaft und dient beispielsweise zur Naherholung (Spaziergänge, Freizeitsport etc.), Berufstätigen als Arbeitsweg oder generell den Anwohner:innen als Weg für ihre Alltagserledigungen (Einkauf etc.). Die Maßnahmenwirksamkeit wird für die Personen getestet, die den BÜ aus westlicher Richtung queren. Abbildung 8 zeigt Bildaufnahmen des Settings.



Abbildung 8 Bildaufnahmen des Settings. Links: Blick auf den BÜ aus südöstlicher Richtung (nach Silla et al., 2019, S. 129). Rechts: Blick von oben auf den BÜ; eingezeichnet sind die Positionen von Blinklicht (roter Kreis) und Markierung (gelbes Rechteck) (Google Maps, 2020).

3.2. Strichprobengewinnung und -beschreibung

Die Stichprobengewinnung innerhalb der Studie wird in Abbildung 9 verdeutlicht.

Phasen\Wochentage	Mo	Di	Mi	Do	Fr	Sa	So
Vergleichsphase ohne Maßnahme Gewöhnung an Verkehrs- überwachungssystem		24 Vpn					
Testphase Markierung		24 Vpn		24 Vpn			
Erholungsphase ohne Maßnahme							
Testphase Blinklicht		24 Vpn		24 Vpn			

Testtag							
vormittags				nachmittags			
3 Vpn	3 Vpn	3 Vpn	3 Vpn	3 Vpn	3 Vpn	3 Vpn	3 Vpn

Abbildung 9 Stichprobengewinnung innerhalb des Versuchsablaufs. Links: Darstellung der Testphasen und Wochentage, an denen Kontrollblickdaten annotiert werden. Rechts: Darstellung der Stichprobenzusammensetzung eines Testtages.

Innerhalb der beiden Testphasen *Markierung* und *Blinklicht* werden an vier Tagen jeweils 24 Personen kodiert, sodass Kontrollblickdaten von jeweils 96 Testpersonen pro Testbedingung vorliegen. Für die Vergleichsphase ohne Maßnahmen, die vor den Testphasen stattfand, wird das Verhalten von 48 Personen kodiert. Hier wird davon ausgegangen, dass für eine repräsentative Erfassung der Kontrollbedingung eine geringere Stichprobengröße ausreicht, da der normale Zustand am BÜ beibehalten wird. Mögliche Verzerrungen durch Gewöhnungseffekte an neue Maßnahmen entfallen. Darüber hinaus findet eine Quotierung der Stichprobe statt. Hierbei werden „Merkmale bzw. ihre Kombinationen“ vorgegeben und „die Auswahl der Teilnehmer erfolgt – gemäß dieser Vorgabe – gezielt und nicht zufällig“ (Hussy et al., 2010, S. 113). Annotiert werden gleichermaßen viele Männer und Frauen sowie gleichermaßen zu Fuß Gehende und Fahrradfahrende, da über die vorliegende Masterarbeit hinaus im Rahmen der vorangegangenen Auswertung durch Illgen (2020) der Geschlechtereinfluss sowie der Einfluss durch die Fortbewegungsart untersucht wurde. Zudem wurden hauptsächlich Personen vormittags zwischen 7:00 Uhr und 8:30 Uhr (80 Personen) und nachmittags zwischen 15:30 Uhr und 16:45 Uhr (82 Personen) beobachtet, da in diesen Zeiträumen die Querungsfrequenz beispielsweise durch Berufstätige auf dem Weg zur und von der Arbeit höher ist als zu anderen Tageszeiten. Mittags zwischen 11:30 Uhr und 12:00 Uhr wurden 40 Personen und abends nach 16:45 Uhr wurden 38 Personen kodiert. Vormittags zwischen 8:30 Uhr und 11:30 Uhr sowie nachmittags zwischen 12:00 Uhr und 15:30 Uhr war die Videoqualität darüber hinaus durch die Sonneneinstrahlung bzw. Beleuchtung des Settings derart eingeschränkt, dass eine eindeutige Kopfbewegung nach links und/ oder rechts zur Identifizierung eines Kontrollblickverhaltens nicht gegeben war.

Die Stichprobe besteht aus 240 beobachteten Personen, die alle in die weitere Analyse mit eingehen. Davon sind 50.4 % männlich und 49.6 % weiblich. Der Unterschied zur angestrebten Verteilung ist darauf zurückzuführen, dass in der Datenauswertung der vorliegenden Masterarbeit eine ursprünglich als weiblich definiert Person als männlich eingestuft wird. Für die Auswertung der Hypothesen werden dadurch keine messbaren Verzerrungen erwartet. 10.8 % der Personen sind der Gruppe der Jugendlichen (14 bis 17 Jahre) zuzuordnen, 73.8 % der Personen der Gruppe der Erwachsenen (18 bis 65 Jahre) und 15.4 % der Gruppe der Senioren (älter als 65 Jahre). In Ermangelung von Fahrradfahrenden oder zu Fuß Gehenden in den für die Annotation ausgewählten Zeiteinheiten wurden auch Personen mit anderen Fortbewegungsarten annotiert, was jedoch zu keinen signifikanten Verzerrungen führen sollte. Die Verteilung der Fortbewegungsarten stellt sich wie folgt dar: 49.6 % zu Fuß Gehende, 48.3 % Fahrradfahrende, 0.4 % E-Scooter-Nutzende sowie jeweils 0.8 % Skatende und motorisiert Fahrende. 39 Personen befanden sich bei der Querung in Gruppen mit ihnen offensichtlich bekannten Personen. Die Gruppengröße variierte zwischen zwei (bei 32 Personen) und drei Personen (bei 7 Personen), wobei die durchschnittliche Gruppengröße $N_{\text{meanGrp}} = 2.17$ betrug. Die Personen hatten grundsätzlich die Möglichkeit, sich über das Projekt SAFER-LC zu informieren. Informationen zu dem Projekt und zur Studie waren an dem Rechnerschränk des Videoüberwachungssystems sichtbar angebracht.

3.3. Ablauf der beiden Annotationen

Grundsätzlich wurden für beide Annotationen dieselben Vorbereitungen getroffen: Die Literaturrecherche zu dem Verhalten von VRU am BÜ hat noch nicht stattgefunden und bisherige Ergebnisse aus dem Projekt SAFER-LC wurden noch nicht kommuniziert, um eine möglichst objektive Datenauswertung zu gewährleisten. Bei beiden Annotierenden handelt es sich um weibliche Psychologiestudierende, die im Rahmen einer Studien- und Masterarbeit die Videodaten auswerten. Daher kann davon ausgegangen werden, dass die Fähigkeiten und das fachliche Vorwissen in etwa ähnlich sind. Im Vorfeld fand eine Annotationsübung statt, bei der nach den Vorgaben für die auswertungsrelevante Annotation 44 Personen kodiert wurden. Informationen zur Annotationsmethode mit entsprechenden Hinweisen für die Identifizierung des Zielverhaltens wurden ebenfalls zur Verfügung gestellt, um eine möglichst gute Schulung der Annotierenden zu gewährleisten. Nach der Übungsphase der Annotierenden fanden mit der Studienleitung reflektierende Gespräche statt und offene Fragen wurden geklärt. Insgesamt kann von einem ausreichenden Training der beiden Annotierenden bezüglich der Annotationssoftware und der Zielverhaltensidentifizierung ausgegangen werden.

Die erste Annotation der Videodaten fand im Frühjahr 2020 statt. Die Vorgehensweise bei der Stichprobenziehung war es, basierend auf einer Liste mit einer ausbalancierten Durchmischung der Bedingungen und der Tageszeiten jeweils vier Personen zu kodieren, von denen möglichst jeweils die Hälfte männlich und weiblich sowie Fahrrad fahrend und zu Fuß gehend sein sollten. Hierbei wurden die durch das Videoüberwachungssystem zugeordneten Personen-IDs für eine spätere Zuordnung mitnotiert. Die zweite Annotation fand im Herbst 2020 statt. Basierend auf der Liste mit den Personen-IDs sowie der Angabe des zugehörigen Videos konnten dieselben Personen in derselben Reihenfolge wie bei der ersten Annotation herausgesucht und kodiert werden. Die beiden Auswertungen fanden somit unabhängig, objektiv und unter nahezu gleichen Bedingungen statt.

3.4. Datenerhebungsmethode

Das Videoüberwachungssystem gehört zu dem DLR-Testfeld AIM (Application Platform for Intelligent Mobility) und wird beschrieben als „an instrument for detection and assessment of traffic behavior in a complex intersection under real-time conditions“ (Knake-Langhorst, Gimm, Frankiewicz & Köster, 2016, S. 2197). Das System ist nach Silla et al. (2019) semi-mobil und besteht aus einem Betonfundament auf dem ein Mast mit den gewünschten Sensoren angebracht werden kann. Darüber hinaus benötigt das System einen Verarbeitungscomputer und diverse Geräte für den Fernzugriff, die in einem Schrank untergebracht sind. Die Sensoren für die vorliegende Studie bestehen für die Maßnahmenpilotierung nach Silla et al. (2019) im Wesentlichen aus Kameras, deren Bilder durch den Computer automatisch zu Trajektorien umgewandelt werden. Aus den Trajektorien ordnet das System den sich vorbei bewegenden Objekten beispielsweise automatisch eine Klassifizierung (zu Fuß Gehender oder Fahrradfahrender) sowie seine Position und die Geschwindigkeit zu. Darüber hinaus erhalten alle Objekte eine Personen ID, die aus einem Buchstaben und einer dreistelligen Nummer besteht. Ethische Belange bezüglich der Videoaufnahmen und Datenschutzerklärungen wurden im Vorfeld mit der Stadt Braunschweig (Deutschland) besprochen und die Studie wurde behördlich genehmigt.

Die definierte physische Messzone für die Beobachtung beginnt 20 m vor und endet einen Meter nach der Bahnstrecke. Die Beobachtungszone ist unterteilt in die Annäherungszone und die Level Crossing-Zone. Letztere beginnt innerhalb der Messzone einen Meter vor der Bahnstrecke und endet einen Meter nach der Bahnstrecke. Die genaue Einteilung ist in Abbildung 10 nachvollziehbar.



Abbildung 10 Einteilung der Messzone anhand einer Momentaufnahme der Videodaten. A entspricht der Annäherungszone, LC entspricht der Level Crossing-Zone. Die Gesichter der Personen auf dem Bild sind aus Datenschutzgründen zusätzlich unkenntlich gemacht.

Die definierten Kategorien des Zielverhaltens für die Videoannotation in dem Annotationsprogramm *ELAN* (Max Planck, 2020) der vorliegenden Masterarbeit ist die Kopfbewegung (nach links und/ oder rechts) des VRU. Aus diesen Annotationen können die Variablen „differenziertes“ und „adäquates Kontrollblickverhalten“ extrahiert werden, die für die Analyse der Maßnahmeneffektivität, des Gewöhnungs- und Wochentageffekts sowie der Interrater-Reliabilität relevant sind.

3.5. Statistische Auswertung

Die Datenauswertung wird anhand der Statistikprogramme *IBM SPSS Statistics Version 20* (*IBM SPSS Statistics for Windows*, 2011) und *R* (*R*, 2010) sowie mit dem Tabellenkalkulationsprogramm *Microsoft Excel* (*Microsoft Office Home and Business 2016*, 2016) im Anschluss an die Videoannotation durchgeführt. Die statistischen Verfahren für die Hypothesentestungen werden im Folgenden vorgestellt. Um bei den multiplen Berechnungen anhand eines Datensatzes einer Alphafehlerkumulierung vorzubeugen, wurde das ermittelte Signifikanzniveau mit Hilfe der Holm-Bonferroni-Korrektur angepasst. Die übliche Bonferroni-Korrektur ist nach Abdi (2010) tendenziell konservativ, was sich auf die Power (Trennschärfe) auswirkt: Das Risiko eines falsch-positiv signifikanten Ergebnis sinkt, wohingegen das Risiko eines falsch-negativen Ergebnisses steigt. „Holm’s procedure is a sequential approach whose goal is to increase the power of the statistical tests while keeping under control the familywise Type I error“ (Abdi, 2010, S. 5). Es finden zwei Korrekturrechnungen getrennt nach der übergeordneten Hauptfragestellung für die Hypothesen zur Maßnahmeneffektivität (insgesamt

drei Tests) und für die Hypothesen zum Gewöhnungs- und Wochentageffekt (insgesamt vier Tests) statt.

3.5.1. Vergleich des Kontrollblickverhaltens der drei Bedingungen.

Die Prüfungen der Hypothesen zum unterschiedlichen Kontrollblickverhalten der Bedingungen erfolgen sowohl deskriptiv durch den prozentualen Vergleich der Häufigkeiten als auch mit Hilfe des Pearson Chi Quadrat Tests. Der Pearson Chi-Quadrat-Test „is based on the simple idea of comparing the frequencies you observe in certain categories to the frequencies you might expect to get in those categories by chance“ (Field, 2018, S. 1064). Für die Verwendung des Pearson Chi-Quadrat-Tests müssen nach Bortz und Schuster (2010) bestimmte Voraussetzungen gelten: (a) Die erwartete Häufigkeit in jeder Kategorie soll mindestens eins betragen und darf bei höchstens 20 % der Kategorien unter fünf liegen, damit die Teststatistik annäherungsweise einer Chi-Quadrat-Verteilung folgt. Ansonsten wird der Exakte Test nach Fisher verwendet. Diese Voraussetzung wird geprüft. (b) Es liegt eine einfache Zufallsstichprobe vor, bei der die Beobachtungen voneinander unabhängig sind. Diese Voraussetzung kann nicht geprüft werden, da die kodierten Personen nicht dazu befragt wurden, ob sie erstmals ohne Kenntnis der anderen Testbedingungen den BÜ überquert haben. Ebenso ist unkontrollierbar, ob die Personen in mehreren Bedingungen kodiert wurden. Grundsätzlich wird davon ausgegangen, dass die Anzahl der Personen, bei denen etwaige abhängige Beobachtungen entstanden sind, statisch marginale Einflüsse auf die Ergebnisse haben. Zur genaueren Analyse und Einschätzung der Effektstärke wird Cramérs-V berechnet. Weiterhin empfiehlt Field (2018) nach einem Pearson Chi-Quadrat-Test die Berechnung von Odds Ratio, um die Beziehung der Variablen zueinander noch zu quantifizieren.

Die Stichprobengröße beträgt bei dem Vergleich zwischen den Testbedingungen $N_{Schrift+Blinklicht} = 192$ ($N_{Schrift} = 96$, $N_{Blinklicht} = 96$). Die Effektgröße wird auf Grund der hypothetischen Herleitung der Maßnahmenwirksamkeit als *mittelgroß* ($w = .3$) eingestuft (Field, 2018). Eine post hoc Analyse der Trennschärfe (Power) mithilfe der Software G*Power (Faul, Erdfelder, Lang & Buchner, 2007) ergab eine Power von $(1 - \beta) = .986$ ($df = 1$, $N_{Schrift+Blinklicht} = 192$, $w = .3$, $\alpha = .05$), wobei ein $(1 - \beta) > .80$ als Mindestmaß für eine zufriedenstellende Power gilt (Field, 2018). Somit ist die Größe der Stichprobe zufriedenstellend gewählt worden. Bei den Vergleichen zwischen der Vergleichsbedingung und den Testbedingungen beträgt die Stichprobengröße $N_{Baseline+Schrift}$ sowie $N_{Baseline+Blinklicht} = 144$ ($N_{Baseline} = 48$). Eine post hoc Analyse der Trennschärfe (Power) mithilfe der Software G*Power (Faul et al., 2007) ergab eine Power von $(1 - \beta) = .95$ ($df = 1$, $N = 144$, $w = .3$, $\alpha = .05$), wobei ein $(1 - \beta) > .80$

als Mindestmaß für eine zufriedenstellende Power gilt (Field, 2018). Somit ist die Größe der Stichprobe zufriedenstellend gewählt worden.

3.5.2. Auftreten des Gewöhnungseffekts.

Die Hypothese, dass der Gewöhnungseffekt bei den Maßnahmen auftritt, wird anhand des Vergleichs des Kontrollblickverhaltens zwischen dem ersten Messtag (Dienstag der ersten Woche) und dem letzten Messtag (Donnerstag der zweiten Woche) entsprechend der Vorgehensweise bei den Vergleichen der drei Bedingungen geprüft. Für die Untersuchungen werden zunächst die Voraussetzungen bezüglich der erwarteten Häufigkeiten für den Pearson Chi-Quadrat-Test geprüft, dann der Pearson Chi-Quadrat-Test oder der Exakte Test nach Fisher durchgeführt, anschließend der χ^2 -Wert in Cramérs-V transformiert sowie Odds Ratio zu dessen Quantifizierung berechnet. Um zu verhindern, dass ein gefundener Gewöhnungseffekt bei dem Vergleich eines Dienstags mit einem Donnerstag durch einen Wochentageffekt zustande gekommen ist oder ein nicht-gefundener Gewöhnungseffekt durch einen Wochentageffekt überlagert wird, wird zusätzlich geprüft, ob der Wochentag einen Einfluss auf die Maßnahmenwirksamkeit hat.

Die Stichprobengröße beträgt bei den Vergleichen zwischen dem ersten Messtag (Dienstag der ersten Woche) und dem letzten Messtag (Donnerstag der zweiten Woche) der beiden Testbedingungen $N_{zweiMesstage} = 48$ ($N_{einMesstag} = 24$). Die Effektgröße wird auf Grund der theoretischen Herleitung zum Vorhandensein des Gewöhnungseffektes als groß ($w = .5$) eingestuft (Field, 2018). Eine post hoc Analyse der Trennschärfe (Power) mithilfe der Software G*Power (Faul et al., 2007) ergab eine Power von $(1 - \beta) = .933$ ($df = 1$, $N_{Messtag} = 48$, $w = .5$, $\alpha = .05$). Somit ist die Größe der Stichprobe für diesen Vergleich entsprechend der Definition von Field (2018) zufriedenstellend gewählt worden.

Für die Berechnung möglicher überlagernder Wochentageffekte beträgt die Stichprobengröße der beiden Testbedingungen jeweils $N_{Schrift} = 96$ und $N_{Blinklicht} = 96$, wobei sich jeweils die Hälfte der Stichprobe auf die beiden Dienstage und die beiden Donnerstage der Testbedingungen verteilt. Die Effektgröße wird als *mittelgroß* ($w = .3$) eingestuft (Field, 2018). Eine post hoc Analyse der Trennschärfe (Power) mithilfe der Software G*Power (Faul et al., 2007) ergab eine Power von $(1 - \beta) = .836$ ($df = 1$, $N_{Schrift} = 48$ und $N_{Blinklicht} = 48$, $w = .5$, $\alpha = .05$). Somit ist die Größe der Stichprobe für diesen Vergleich entsprechend der Definition von Field (2018) zufriedenstellend gewählt worden.

3.5.3. Bestimmung der Interrater-Reliabilität.

Ein gängiges Verfahren, „mit dem man die Übereinstimmung von zwei Klassifikationen derselben Objekte erfassen und überprüfen kann“, ist das Kappa-Maß (Bortz & Schuster, 2010, S. 582). Die Maßzahl Cohens K bietet für die Beurteilungen numerisch skaliertter Variablen die Möglichkeit, die Beurteilerübereinstimmung, „die über das Ausmaß an durch Zufall zu erwartender Übereinstimmung hinaus geht, in standardisierter Form zu quantifizieren und zu evaluieren“ (Wirtz & Kutschmann, 2007, S. 7). Problematisch bei der Interpretation von Cohens K ist, dass „insbesondere die Häufigkeit, mit der ein Merkmal tatsächlich vorliegt (Prävalenz), die Koeffizientenausprägung beeinflusst: Je stärker die Grundrate eines Merkmals von 50 % abweicht, desto niedriger ist K unter ansonsten gleichen Umständen ausgeprägt“ (Wirtz & Kutschmann, 2007, S. 5). Wirtz und Kutschmann (2007) schlagen deshalb vor, sofern die Variablen dichotom ausgeprägt sind, vorzugsweise Yules Y zu verwenden, was in der vorliegenden Masterarbeit umgesetzt wird. Yules Y basiert auf dem Odds Ratio und erlaubt im direkten Vergleich mit Cohens K auch eine Schätzung darüber, wie hoch Cohens K durch die Grundrate des Merkmals beeinflusst wird. Zusätzlich wird anhand der prozentualen Häufigkeiten der Übereinstimmungen geschaut, ob es systematische Zusammenhänge zwischen der Übereinstimmungsquote und den erhobenen Variablen gibt.

4. Ergebnisse

Im Folgenden werden die Ergebnisse der statistischen Analysen zur Maßnahmeneffektivität, zum Gewöhnungs- und Wochentageffekt sowie zur Interrater-Reliabilität entsprechend der aufgestellten Hypothesen vorgestellt.

4.1. Maßnahmeneffektivität

Das differenzierte Kontrollblickverhalten der drei Testbedingungen ist in Abbildung 11 dargestellt und tabellarisch in Anhang A aufgeführt. In der Abbildung sind die prozentualen Unterschiede zwischen den Testbedingungen bezüglich der drei kategorialen Ausprägungen des Kontrollblicks (gar nicht, in eine Richtung und in beide Richtungen) dargestellt.

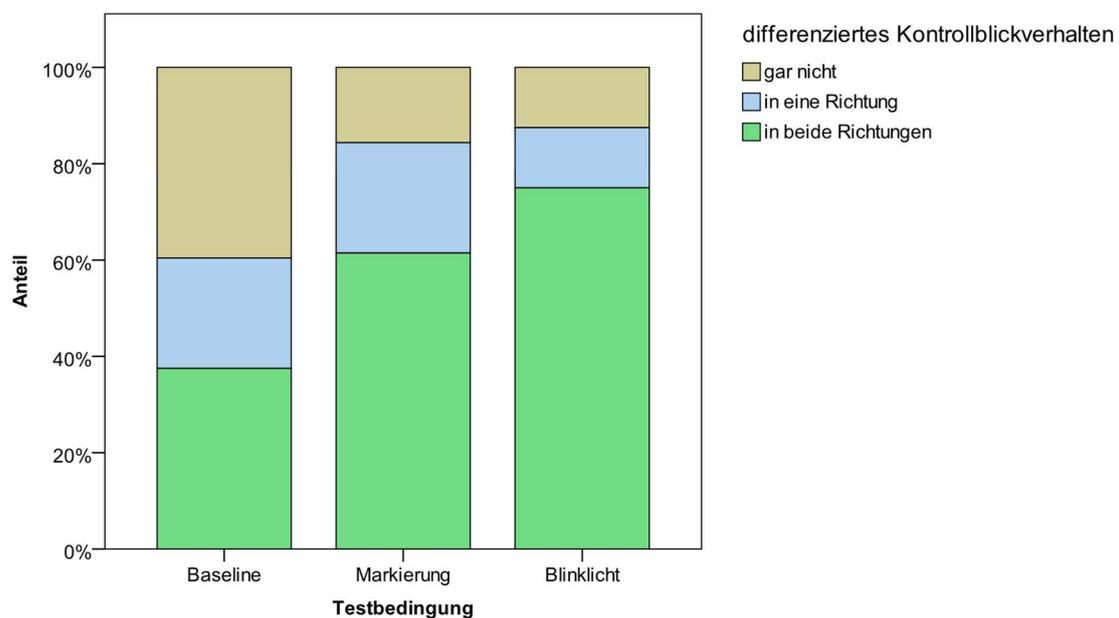


Abbildung 11 Differenzierte Darstellung des Kontrollblickverhaltens innerhalb der Testbedingungen.

In dieser Darstellung ist ein Unterschied zwischen dem adäquaten Kontrollblickverhalten der Kontrollphase (*Baseline*) und den Testphasen mit den Maßnahmen (*Markierung* und *Blinklicht*) erkennbar: In beide Richtungen schauten in der Kontrollphase 37.5 %, wohingegen bei der Testphase mit Markierung 61.5 % und bei der Testphase mit Blinklicht 75 % der beobachteten Personen ein adäquates Kontrollblickverhalten zeigten. Der Anteil der Personen, die in nur eine Richtung schauen, ist bei den Testbedingungen *Baseline* und *Markierung* exakt gleich hoch (22.9 %). Evident ist außerdem, dass in der Vergleichsphase der Anteil der Personen, die gar nicht schauen, höher ist als in den Testphasen (39.6 % vs. 15.5 % bei der Markierung und 12.5 % bei dem Blinklicht). Bei dem prozentualen Vergleich des Kontrollblickverhaltens zwischen den Testphasen *Markierung* und *Blinklicht* ist ersichtlich,

dass in etwa gleich viele Personen gar keinen Kontrollblick durchgeführt haben (15.6 % bei der Markierung und 12.5 % bei dem Blinklicht), es aber Unterschiede bei dem Blick in eine Richtung gegeben hat (22.9 % bei der Markierung und 12.5 % bei dem Blinklicht). Außerdem ist ein Unterschied zwischen dem Kontrollblick in beide Richtungen zwischen den Testphasen mit den Maßnahmen aufgetreten: 14.5 % der beobachteten Personen schauten bei der Testphase mit Markierung weniger in beide Richtungen als bei der Testphase mit Blinklicht.

Der Fokus der vorliegenden Studie liegt auf der Untersuchung der Maßnahmenwirksamkeit bezüglich des korrekten Kontrollblickverhaltens, das entsprechend der Anforderungen der Straßenverkehrsordnung in beide Richtungen stattfinden soll. In Tabelle 1 sind die Häufigkeiten der beiden Ausprägungen inadäquater und adäquater Kontrollblick entsprechend der drei Testbedingungen aufgetragen. Die Einzelvergleiche mit den Chi-Quadrat-Tests finden entsprechend der Hypothesen statt. Sie sind in Anhang B detailliert tabellarisch aufgeführt und werden im Folgenden vorgestellt.

Tabelle 1

Häufigkeiten des adäquaten Kontrollblickverhaltens innerhalb der Testbedingungen

Testbedingung	Kontrollblick [% innerhalb der Testbedingung]		N _{Bedingung}
	inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen	
Baseline	62.5	37.5	48
Markierung	38.5	61.5	96
Blinklicht	25.0	75.0	96

Ein Chi-Quadrat-Test wurde zwischen dem Kontrollblickverhalten und den Testbedingungen *Baseline* und *Markierung* durchgeführt. Keine erwarteten Zelloberhäufigkeiten waren kleiner als fünf. Es gab einen statistisch hoch signifikanten Zusammenhang zwischen dem Kontrollblick und den Testbedingungen *Baseline* und *Markierung*, $\chi^2(1) = 7.383$, $p = .008$, $Cramérs-V = .226$. Die Effektstärke wird nach Field (2018) als klein bewertet. Die Wahrscheinlichkeit, ein adäquates Kontrollblickverhalten zu zeigen, erhöht sich durch die Markierung um den Faktor $OR = 2.685$ (95 % CI: 1.301, 5.43).

Ein Chi-Quadrat-Test wurde zwischen dem Kontrollblickverhalten und den Testbedingungen *Baseline* und *Blinklicht* durchgeführt. Keine erwarteten Zelloberhäufigkeiten waren kleiner als fünf. Es gab einen statistisch höchst signifikanten Zusammenhang zwischen dem

Kontrollblick und den Testbedingungen, $\chi^2(1) = 19.2$, $p < .001$, $Cramérs-V = .365$. Die Effektstärke wird nach Field (2018) als *mittelstark* bewertet. Die Wahrscheinlichkeit, ein adäquates Kontrollblickverhalten zu zeigen, erhöht sich durch das Blinklicht um den Faktor $OR = 5$ (95 % CI: 2.739, 10.5312).

Ein Chi-Quadrat-Test wurde zwischen dem Kontrollblickverhalten und den Testbedingungen *Markierung* und *Blinklicht* durchgeführt. Keine erwarteten Zelhäufigkeiten waren kleiner als fünf. Es gab einen statistisch signifikanten Zusammenhang zwischen dem Kontrollblick und den Testbedingungen, $\chi^2(1) = 4.061$, $p = .022$, $Cramérs-V = .145$. Die Effektstärke wird nach Field (2018) als *klein* bewertet. Die Wahrscheinlichkeit, ein adäquates Kontrollblickverhalten zu zeigen, erhöht sich durch das Blinklicht um den Faktor $OR = 1.881$ (95 % CI: 1.014, 3.492).

4.2. Gewöhnungs- und Wochentageffekt

In Abbildung 12 sind die acht Messtage der Testbedingungen mit Maßnahmen sowie die zugehörigen, prozentualen Häufigkeiten des adäquaten Kontrollblickverhaltens abgebildet. Aus dem prozentualen Vergleich der Häufigkeiten ist ersichtlich, ob ein Gewöhnungseffekt zwischen dem jeweils ersten und letzten Messtag der Testphasen mit Maßnahmen aufgetreten sein könnte. In Anhang C sind die Daten tabellarisch einsehbar. Zusätzlich wird untersucht, ob die Wochentage (Dienstag und Donnerstag) mögliche Gewöhnungseffekte überdecken könnte. Darüber hinaus wird bei der Betrachtung der prozentualen Häufigkeiten in der Abbildung 12 deutlich, dass die Testbedingungen in sich eine annähernd homogene Verteilung des adäquaten Kontrollblickverhaltens an den verschiedenen Messtagen aufzeigen.

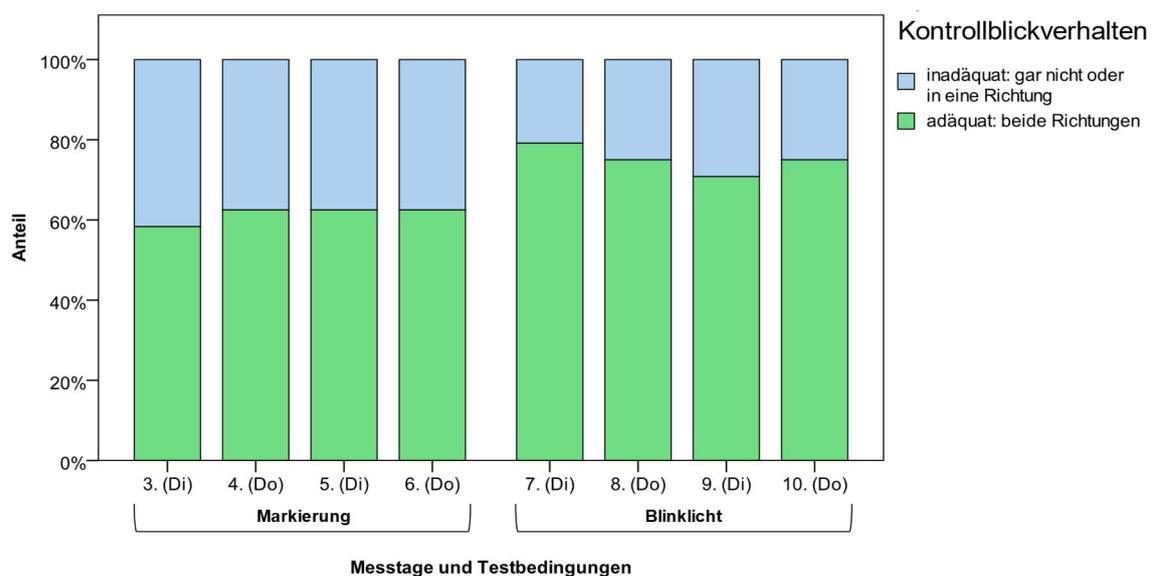


Abbildung 12 Prozentuales Auftreten adäquaten Kontrollblickverhaltens an den acht Messtagen der Testbedingungen mit Maßnahmen.

In der Testphase mit der Markierung schauten an dem ersten und dem letzten Messtag der Testbedingung nahezu gleich viele Personen in beide Richtungen nach einem Schienenfahrzeug (dritter Messtag 58.3 %, sechster Messtag 62.5 %). In der Testphase mit dem Blinklicht schauten an dem ersten und dem letzten Messtag der Testbedingung ebenfalls nahezu gleich viele Personen in beide Richtungen nach einem Schienenfahrzeug (79.2 % am siebten Messtag, 75 % am zehnten Messtag). Bei dem Vergleich der Messtage der Testbedingungen sind augenscheinlich keine Effekte durch die Wochentage ersichtlich: Bei der Testbedingung mit der Markierung werden hierbei die Dienstage (dritter und fünfter Messtag) mit den Donnerstagen (vierter und sechster Messtag) nebeneinandergestellt. An den Dienstagen schauten mit 60.4 % nahezu gleich viele Personen in beide Richtungen wie an den Donnerstagen mit 62.5 %. Bei der Testbedingung mit dem Blinklicht werden gleichermaßen die Dienstage (siebter und neuntes Messtag) mit den Donnerstagen (achter und zehnter Messtag) nebeneinandergestellt. An den beiden Tagen schauten gleich viele Personen in beide Richtungen (75 %).

Um statistisch gesichert einen Gewöhnungseffekt an die Maßnahmen festzustellen, wurden Chi-Quadrat-Tests zwischen dem jeweiligen Kontrollblickverhalten der beiden Testbedingungen und dem ersten und letzten Messtag durchgeführt. Die Ergebnisse sind detailliert in Anhang D dargestellt. Keine erwarteten Zellhäufigkeiten waren kleiner als fünf, womit die Voraussetzung für den Pearson Chi-Quadrat-Test als erfüllt gilt. Es gab sowohl für die Testphase mit Markierung ($\chi^2(1) = .087, p = 1.$) als auch für die mit Blinklicht ($\chi^2(1) = .118, p = 1.$) keinen statistisch signifikanten Zusammenhang zwischen dem Kontrollblickverhalten und den getesteten Tagen.

Um einen Effekt der Wochentage auf die Wirksamkeit der Maßnahmen statistisch festzustellen, wurden Chi-Quadrat-Tests zwischen dem Kontrollblickverhalten und den getesteten Wochentagen (jeweils zwei Dienstage und Donnerstage) der beiden Testbedingungen durchgeführt. Die Ergebnisse sind detailliert in Anhang E dargestellt. Keine erwarteten Zellhäufigkeiten waren kleiner als fünf, womit die Voraussetzung für den Pearson Chi-Quadrat-Test für beide Bedingungen als erfüllt gilt. Es gab sowohl für die Testphase mit Markierung ($\chi^2(1) = .044, p = 1.$) als auch für die mit Blinklicht ($\chi^2(1) = .0, p = 1.$) keinen statistisch signifikanten Zusammenhang zwischen dem Kontrollblickverhalten und den getesteten Wochentagen.

4.3. Interrater-Reliabilität

Um festzustellen, ob die Beurteilungen zu den Eigenschaften der VRU (Geschlecht, Altersgruppe, Fortbewegungsmittel) der beiden Annotierenden mindestens *mittelmäßig* nach Landis

und Koch (1977) übereinstimmen, wurde die Beurteilungsübereinstimmung anhand von Cohens K bestimmt. Bei der Variable *Geschlecht* konnte zusätzlich Yules Y berechnet werden, da selbige dichotom ausgeprägt ist. Die Ergebnisse sind in Tabelle 2 aufgeführt. Die entsprechenden Häufigkeitstabellen sind in Anhang F einsehbar.

Tabelle 2

Analyse der Interrater-Reliabilität zu den Eigenschaften der VRU

	Gültige Fälle	Cohens K	asympt. SE	95 % Konfidenzintervall		OR ^a	Yules Y ^a	K Beurteilung ^b
				Oberer Grenzwert	Unterer Grenzwert			
Geschlecht	239	.916	.026	.917	.915	524.4	.916	nahezu perfekt
Altersgruppe	240	.517	.064	.520	.514	-	-	mittelmäßig
VRU-Typ	240	.944	.021	.945	.943	-	-	nahezu perfekt

Anmerkungen. VRU-Typ = vulnerable road user - Typ (entspricht dem Fortbewegungsmittel der Versuchsperson). Vergleich der Beurteilungen von zwei Annotierenden. Zwei Ausprägungen des Geschlechts: männlich, weiblich. Drei Abstufungen der Altersgruppe: Jugendliche, Erwachsene, Senioren. Drei VRU-Typen: zu Fuß Gehende, Fahrradfahrende, andere Fortbewegungsmittel.

^a Yules Y basiert auf OR und ist lediglich bei dem Vergleich von zwei dichotom ausgeprägten Variablen berechenbar.

^b Beurteilung von K entsprechend der Kategorisierung von Landis und Koch (1977).

Die Übereinstimmungen können nach Landis und Koch (1977) als *nahezu perfekt* für die Variablen *Geschlecht* (Yules Y = .916, OR = 524.4) und *Fortbewegungsmittel* (K = .944, 95 % CI: .943, .945) sowie als *mittelmäßig* für die Variable *Altersgruppe* (K = .517, 95 % CI: .514, .520) bewertet werden. Yules Y stimmt für die Variable *Geschlecht* mit Cohens K überein (K = .616, 95 % CI: .915, .917).

Für die Analyse, ob die Urteile zu dem Kontrollblickverhalten der beiden Annotierenden mindestens *mittelmäßig* übereinstimmen, wurde die Interrater-Reliabilität anhand von Cohens K bestimmt. Um eine differenziertere Betrachtung der Übereinstimmung zu ermöglichen, wurde über die Betrachtung des adäquaten Kontrollblickverhaltens mit seinen zwei Ausprägungen hinaus ein Vergleich der Beurteilungen zu den drei differenzierten Kategorien des Kontrollblicks in gar keine, Blick in eine und Blick in beide Richtungen durchgeführt. Bei der Analyse des adäquaten Kontrollblickverhaltens konnte zusätzlich Yules Y berechnet werden, da die Variable dichotom ausgeprägt ist. Die Ergebnisse sind in Tabelle 3 aufgeführt. Die entsprechenden Häufigkeitstabellen sind detailliert in Anhang G einsehbar.

Tabelle 3

Analyse der Interrater-Reliabilität zu dem beobachteten Kontrollblickverhalten der VRU

	Gültige Fälle	Cohens K	asympt. SE	95 % Konfidenzintervall		OR ^a	Yules Y ^a	K Beurteilung ^b
				Oberer Grenzwert	Unterer Grenzwert			
adäquates KBV	240	.601	.052	.604	.599	17.526	.614	substanziell
differenz. KBV	240	.502	.047	.504	.500	-	-	mittelmäßig

Anmerkungen. KBV = Kontrollblickverhalten. Vergleich der Beurteilungen von zwei Annotierenden. Zwei Ausprägungen des adäquaten KBV: adäquat in beide Richtungen, inadäquat in gar keine oder in eine Richtung. Drei Abstufungen des differenzierten KBV: Blick in gar keine Richtung, Blick in eine Richtung, Blick in beide Richtungen.

^a OR und Yules Y sind lediglich bei dem Vergleich von zwei dichotom ausgeprägten Variablen berechenbar.

^b Beurteilung von K entsprechend der Kategorisierung von Landis und Koch (1977).

Die Übereinstimmungen können nach Landis und Koch (1977) als *substanziell* für die Variable des *adäquaten Kontrollblickverhaltens* (Yules $Y = .614$, $OR = 17.526$) sowie als *mittelmäßig* für das *differenzierte Kontrollblickverhalten* ($K = .502$, 95 % CI: .500, .504) bewertet werden. Yules Y ist für die Variable des *adäquaten Kontrollblickverhaltens* im Vergleich mit Cohens K nahezu gleich groß ($K = .601$, 95 % CI: .599, .604).

5. Diskussion

Die Ergebnisse der Studie zur Förderung von adäquatem Kontrollblickverhalten der VRU am nicht-technisch gesicherten BÜ durch zwei kostengünstige, innovative Maßnahmen hat gezeigt, dass sich im Vergleich zu der Bedingung ohne Maßnahmen die Wahrscheinlichkeit für einen sicheren Kontrollblick in beide Richtungen erhöht hat. Die Ergebnisse deuten darauf hin, dass die Maßnahme mit dem Blinklicht, das bei Annäherung aktiviert wird, eine höhere Wirksamkeit im Vergleich mit der Maßnahme mit der Straßenmarkierung *Kommt ein Zug?* aufweist. Ein Gewöhnungs- sowie ein eventuell überlagernder Wochentageffekt, wurden nicht gefunden. Die Analyse der Interrater-Reliabilitäten für relevante Variablen bezüglich des Kontrollblickverhaltens und verschiedener Personenvariablen bestätigt, dass sich die Methode der Videoannotation anhand von Aufnahmen eines Verkehrsüberwachungssystems hinsichtlich des Gütekriteriums der Objektivität für die Analyse des Blickverhaltens von Personen unter den in der Studie vorgegebenen Bedingungen eignet.

Im folgenden Kapitel werden die Ergebnisse zur Maßnahmeneffektivität, zum Gewöhnungs- und Wochentageffekt sowie zur Interrater-Reliabilität genauer beschrieben und interpretiert. Basierend auf den Studienergebnissen werden anschließend Implikationen für die Praxis diskutiert, die Empfehlungen für die Gestaltung nicht-technisch gesicherter BÜ beinhalten. Hierbei steht im Vordergrund, dass entsprechend der Forderung von Dressler und Silla, et al. (2020) sowohl Wahrnehmung, Aufmerksamkeits- und Bewertungsprozesse als auch die Motivation seitens der VRU berücksichtigt werden müssen, wenn durch die Umweltgestaltung sicherheitsrelevante Verhaltensweisen gefördert werden sollen. Abschließend werden Limitationen der Forschungsmethode der vorliegenden Studie erläutert und offene Forschungsfragen besprochen, welche zukünftige Untersuchungen anregen können.

5.1. Maßnahmeneffektivität

Die Ergebnisse zeigen deutliche Unterschiede zwischen den drei Testbedingungen hinsichtlich des adäquaten Kontrollblickverhaltens von Straßennutzenden vor dem BÜ. Durch die Straßenmarkierung erhöhte sich die Auftretenswahrscheinlichkeit eines adäquaten Kontrollblickverhaltens bei den beobachteten Personen im Vergleich mit der Kontrollbedingung ohne Maßnahmen signifikant von 37.5 % auf 61.5 %. Die beobachtete Effektivität des Blinklichts war noch größer: Mit 75.5 % schauten nahezu doppelt so viele Personen adäquat in beide Richtungen nach einem nahenden Schienenfahrzeug wie in der Bedingung ohne Maßnahmen. Dementsprechend verringerten sich die Anteile der Personen, die entweder nur in eine Rich-

tung oder gar nicht schauen. Dies bestätigt insgesamt die Annahme, dass die beiden getesteten Maßnahmen dazu in der Lage sind, in ihrer Funktion als Prompt und als Warnung dafür zu sorgen, dass mehr Personen zu adaptivem, sicherem Verhalten angeregt werden. Mögliche Erklärungen für die Wirksamkeit der Maßnahmen bieten Modelle der kognitiven Informationsverarbeitung. Basierend auf der Annahme von Zachrisson und Boks (2010) könnten die Maßnahmen innovativ genug gewesen sein, um die Interaktion zwischen Menschen und Infrastruktur derart zu beeinflussen, dass bei einem Teil der Querenden ungewünschte teilautomatische Verhaltensweisen verhindert wurden. Das Ziel von Maßnahmen zur Erhöhung der Sicherheit an nicht-technisch gesicherten BÜ sollte nach Gripenkoven und Dietsch (2015) sowie Rudin-Brown et al. (2014) sein, tiefere Verarbeitungsschritte zu aktivieren, die über das bloße Erkennen des BÜ als solchen hinausgehen. Diese Verarbeitung sollte entweder zu einer Aktivierung vorhandener Schemata und Skripte adäquaten Verhaltens oder bei Nichtvorhandensein und bei fehlerhaften Skripten (z.B. ein „Kein-Zug-Schema“, wenn in der Vergangenheit nie ein Schienenfahrzeug in der Situation wahrgenommen wurde) zur aktiven Auseinandersetzung mit der Umgebung führen, wodurch idealerweise adäquate Verhaltensstrategien durch die Personen zur Bewältigung entwickelt werden. Eine weitere Ursache für die Verbesserung des Kontrollblickverhaltens könnte sein, dass durch die zusätzliche Warnbotschaft im Bereich des nicht-technisch gesicherten BÜ die Bedrohungseinschätzung entsprechend der PMT von Rogers et al. (1983) erhöht wurde, was zu einer höheren Schutzmotivation und somit zu einer höheren Intention zur Ausführung adäquaten Verhaltens geführt hat. Die Gestaltung der Information scheint entsprechend der Attribute für gute Warnbotschaften von Conzola und Wogalter (2001) grundsätzlich gelungen zu sein. Auffallend bei der Betrachtung der Häufigkeiten des differenzierten Kontrollblickverhaltens in seinen drei Ausprägungen (gar nicht, in eine Richtung, in beide Richtungen) ist, dass sich die Anteile der Personen, die gar nicht schauen, zwischen den beiden Testbedingungen mit Maßnahmen nicht unterscheiden. Der Grund dafür könnte beispielsweise sein, dass die Personen zu stark abgelenkt waren, um die Prompts oder die Warnbotschaften wahrzunehmen und im Rahmen tiefgreifenderer Verarbeitungen zu verstehen (z.B. durch starke interne Ablenkungen wie Grübeln oder externe Ablenkungen durch Gespräche mit Begleitpersonen). Eine weitere Erklärungsmöglichkeit ist, dass die Motivation für inadäquates Verhalten im Straßenverkehr bei einem gewissen Anteil der Personen zu hoch ist und somit bewusste Verstöße entsprechend der Fehlerklassifikation von Reason (1990) in Kauf genommen werden. Hierbei geschehen willentlich Taten, die zu unsicheren Handlungen führen (z.B. ein grundsätzlich rasanter Fahrstil beim Fahrradfahren, bei dem es keine Gelegenheit für adäquates Kontrollblickverhalten gibt).

Die Implementierung des Blinklichts führte zu einer signifikanten Erhöhung der Auftretenswahrscheinlichkeit von adäquatem Kontrollblickverhalten im Vergleich zu der Bedingung mit der Straßenmarkierung. Der Grund für die größere Maßnahmeneffektivität könnte die höhere Salienz des Blinklichts sein. Die Eigenschaften von Stimuli wie Farbe, Größe und Bewegung, die die visuelle Aufmerksamkeit *unbestreitbar effektiv* nach Wolfe und Horowitz (2004) lenken können, legen die Vorteile seitens des Blinklichts nahe. Die Straßenmarkierung ist auf dem Boden außerhalb der Hauptblickrichtung angebracht, wenn sich Personen in eine Richtung bewegen. Darüber hinaus verfügt das Blinklicht über das zusätzliche Attribut des Blinkens. Weiterhin ist denkbar, dass ein Teil der beobachteten Personen die Schrift auf dem Boden nicht verstanden haben könnte, da es sich entweder um Analphabeten handelte oder um Personen, die kein Deutsch verstehen. 2018 befanden sich nach Grotlüschen und Buddeberg (2020) zwei Millionen der deutsch sprechenden Erwachsenen (4 %) auf einem Leseneveu, bei dem die Personen nicht in der Lage sind, Sätze zu verstehen. Kritisch anzumerken ist, dass bei dem Blinklicht theoretisch das Problem der Fehlinterpretation besteht: Querende könnten das Blinken als konkrete Warnung vor einem nahenden Zug verstehen wie es bei technisch-gesicherten BÜ der Fall ist, bei denen eine rot blinkende Signalanlage über dem Andreaskreuz ein kreuzendes Schienenfahrzeug ankündigt. Sofern solche Fehlinterpretationen auftreten, könnte dies zu Einbußen bezüglich der Vertrauenswürdigkeit der Quelle führen, was nachhaltige Folgen für die Compliance bezüglich derartiger Maßnahmen hätte. Vielleicht nehmen die Personen zukünftig entsprechende Warnungen nicht mehr ernst, was Unfälle begünstigen würde. Bei der Annotation der Videoaufnahmen für die vorliegende Masterarbeit ist jedoch nicht aufgefallen, dass Personen in der Testphase mit dem Blinklicht an der Bahnstrecke stehen geblieben sind oder ähnliche Verhaltensmuster zeigten, die ein derartiges Missverstehen nahelegen. Wenn Querende fälschlicherweise davon ausgehen, dass ein Zug käme, wäre eine Folge sicherlich die Reduzierung der Geschwindigkeit in der Annäherungszone. Bei früheren Auswertungen der Geschwindigkeitsdaten der Studie von Silla et al. (2019) ist zudem nicht nachgewiesen worden, dass es Unterschiede bei der Geschwindigkeit in der Annäherungszone des BÜ der Straßenverkehrsteilnehmenden zwischen den Bedingungen gab.

5.2. Gewöhnungs- und Wochentageffekt

Die Hypothese, dass über den Erprobungszeitraum eine Gewöhnung an die implementierten Maßnahmen auftreten würde, ist anhand der vorliegenden Daten nicht bestätigt worden. Ein Wochentageffekt, der einen möglichen Gewöhnungseffekt verschleiern könnte, konnte

ebenfalls nicht nachgewiesen werden. Somit liefert die Untersuchung zunächst keine Hinweise auf ein Nachlassen der Wirkung über die ersten zwei Wochen der Implementierung einer Maßnahme. Da jedoch keine Daten über die tatsächliche Erfahrung der beobachteten Personen mit der jeweiligen Maßnahme vorlagen, ist unklar, ob die zu Ende des jeweiligen Testzeitraums beobachteten Personen in der Tat bereits häufiger mit der Maßnahme konfrontiert waren als die zu Anfang der Testphase beobachteten. Zudem waren die Testphasen mit zwei Wochen zu kurz, um tragfähige Aussagen über mögliche Langzeiteffekte zu erlauben. Somit kann bislang nicht ausgeschlossen werden, dass sich die Wirksamkeit der Maßnahmen über die Zeit verändert.

5.3. Interrater-Reliabilität

Die Hypothesen zur Interrater-Reliabilität konnten bestätigt werden. Mittels Cohens K konnten die erhaltenen Übereinstimmungswerte entsprechend der Einstufung von Landis und Koch (1977) für alle betrachteten Variablen mindestens als *mittelmäßig* ($K > .41$) eingestuft werden. Besonders hohe Übereinstimmungen zeigten sich bei dem Geschlecht und dem Fortbewegungsmittel der beobachteten Personen (*nahezu perfekt*). Die Übereinstimmung bei der Beurteilung der Altersgruppe konnte als *substanziell* beurteilt werden. Eine nähere Betrachtung der deskriptiven Statistik zeigt, dass die Unterschiede bei der Alterseinschätzung vor allem im Grenzbereich zwischen Jugendlichen und Erwachsenen sowie Erwachsenen und Senior:innen auftraten. Die unterschiedlichen Einschätzungen, ob eine Person knapp ober- oder unterhalb einer festen Altersgrenze einzuordnen ist, geschahen im Rahmen des erwartbaren Spielraums bei der subjektiven Einschätzung des Alters von Personen (siehe Anhang F2). Für das Kontrollblickverhalten konnte bei der dichotomen Einschätzung zum adäquaten Kontrollblickverhalten (in beide Richtungen vs. gar nicht/ in eine Richtung) die Übereinstimmung als *substanziell* bewertet werden. Bei der Betrachtung der drei differenzierteren Kategorien des Kontrollblickverhaltens (gar nicht, in eine Richtung und in beide Richtungen) ist die Übereinstimmung *mittelmäßig*. Eine nähere Analyse durch den Vergleich der prozentualen Häufigkeiten ergab, dass die zweite Annotation das Kontrollblickverhalten etwas liberaler im Sinne von toleranter und kulanter beurteilte. Bei dem Vergleich der Geschlechts- und der Kontrollblickurteile konnte ebenfalls Yules Y berechnet werden, was eine Korrektur des Ausmaßes ermöglicht, in dem Cohens K durch die Grundrate des Merkmals beeinflusst wird. Es weicht unwesentlich von Cohens K ab, womit die Schlussfolgerungen zur Interrater-Reliabilität unterstützt werden.

Somit haben beide Annotierende sowohl das adäquate und differenzierte Kontrollblickverhalten als auch die erfassten Personenvariablen (Geschlecht, Altersgruppe, Fortbewegungsmittel) zufriedenstellend übereinstimmend beurteilt. Schlussfolgernd eignet sich die Methode der Videoannotation mit Hilfe des Verkehrsüberwachungssystems aus dem DLR-Testfeld AIM hinsichtlich des Gütekriteriums der Objektivität grundsätzlich für die Analyse des Blickverhaltens von Personen unter denen in der Studie vorliegenden Bedingungen.

5.4. Implikationen für die Praxis

Die Ergebnisse der vorliegenden Masterarbeit lassen darauf schließen, dass es in jedem Fall sinnvoll ist, die derzeitige Beschilderung in Form des Andreaskreuz um weitere Informationen zur Förderung von sicheren Verhaltensweisen bei der Querung von nicht-technisch gesicherten BÜ zu ergänzen. Zu diesem Fazit kommen auch Rudin-Brown et al. (2014), die die Unfallzahlen an kanadischen BÜ über einen Zeitraum von zehn Jahren analysiert haben: „Advising drivers of the presence of a level crossing is not enough. Warning drivers about the crossing and providing sufficient time for them to detect a train and decide what to do are imperative” (S. 102). Demnach eignen sich die beiden getesteten Maßnahmen direkt in der Annäherungszone des BÜ dazu, vor einer nahenden potenziellen Gefahr zu warnen. Aus den genannten Gründen zu möglichen Fehlschlüssen in Folge des Blinklichts wird jedoch dringend davon abgeraten, selbiges in der getesteten Gestaltungsvariante als alleinige Maßnahme an nicht-technisch gesicherten BÜ zu installieren. Die Ergänzung von Informationen kann die potenzielle Fehlinterpretation, dass zeitnah ein Schienenfahrzeug queren wird, jedoch verhindern. Beispielsweise könnte hinter dem Symbol des Zuges noch ein Fragezeichen folgen oder unterhalb des Blinklichts könnte ein Zusatzschild mit der Aufschrift *Kommt ein Zug?* installiert werden, wie in Abbildung 13 dargestellt ist.



Abbildung 13 Zwei Ideen zur Optimierung des Blinklichts zur Vermeidung von Fehlinterpretationen. a. mit einem Zusatzschild unterhalb des Blinklichts. b. mit einem ergänzenden Fragezeichen hinter dem Zugsymbol (eigene Modifikation einer Abbildung des ursprünglichen Blinklichts aus Silla et al., 2019, S. 127).

Ein weiterer zur Diskussion stehender Punkt hinsichtlich der Maßnahme mit dem Blinklicht ist, dass die langfristigen Kosten genau kalkuliert werden sollten. Für einen flächendeckenden Rollout an nicht-technisch gesicherten BÜ sollten die Kosten für die Installation mit entsprechender Stromversorgung (Anschluss an das lokale Stromnetz oder alternative Energieversorgung beispielsweise anhand von Solarbetrieb) sowie der Aufwand für die Wartung langfristig geplant werden. Wenn bereits ein Stromanschluss vorhanden und auch die Wartung händelbar ist, sollte eine überarbeitete Version des Blinklichts gegenüber der Straßenmarkierung aus Gründen der Salienz bevorzugt werden. Für den Fall, dass eine Stromversorgung und/oder Wartung nicht umsetzbar sein sollten und die Straßenverhältnisse ein Aufbringen der Markierung nicht gewährleisten (z.B. auf unbefestigten Schotterwegen), könnte nach Silla et al. (2019) auch ein Straßenschild mit der Aufschrift *Kommt ein Zug?* installiert werden. Darüber hinaus ist in Anbetracht von Schnee-, Laubfall oder Straßenverschmutzungen eine Umwandlung der Straßenmarkierung in ein Straßenschild zu diskutieren. Zur Erhöhung der Salienz der Straßenmarkierung war eine weitere Idee von Silla et al. (2019) die Schriftfarbe auffallender zu wählen, was in Abbildung 14 exemplarisch umgesetzt wurde: Der Schriftzug könnte beispielsweise in orange-gelber Signalfarbe aufgetragen werden



Abbildung 14 Idee zur Optimierung der Straßenmarkierung zur Erhöhung der Salienz (eigene Modifikation einer Abbildung der ursprünglichen Markierung aus Silla et al., 2019, S. 126).

Die Straßenmarkierung sollte außerdem dahingehend überdacht werden, ob er für die Verbesserung der Verständlichkeit für Personen, die (noch) nicht lesen können, durch zusätzliche Symbole ergänzt wird (vgl. Conzola & Wogalter, 2001). Ein Entwurf dafür ist mit dem bereits beschriebenen Icon eines Zuges in Abbildung 15 dargestellt.



Abbildung 15 Idee zur Optimierung der Straßenmarkierung zur Verbesserung der Verständlichkeit auch für Personen, die (noch) nicht lesen können (eigene Modifikation einer Abbildung der ursprünglichen Markierung aus Silla et al., 2019, S. 126).

5.5. Limitationen der Untersuchungsmethode

Restriktionen bezüglich der Ergebnisse der vorliegenden Masterarbeit ergeben sich bei der Stichprobenziehung, der zeitlichen Skalierung der Studie und der Schulung der Annotierenden.

Eine Kritik an der Studie ist die Ziehung der Stichprobe. Es handelt sich um eine quitierte Stichprobe, die zu den nicht-probabilistischen Stichproben gehört und i.d.R. für inferenzstatistische Auswertung ungeeignet sind: die realistische Generalisierbarkeit der Befunde ist fragwürdig (Bortz & Schuster, 2010; Hussy et al., 2010). Grundsätzlich ist nach Matthies et al. (2011) für die Implementierung und Evaluation die Anpassung der Maßnahme an die real zu erwartende Zielgruppe essenziell, weshalb die Studienergebnisse bezüglich der Übertragbarkeit in jedem Fall hinterfragt werden müssen. In der in Silla et al. (2019) beschriebenen Voruntersuchung der Maßnahmenwirksamkeit anhand von 230 zufällig querenden Personen

wurde bereits festgestellt, dass nicht gleichermaßen viele zu Fuß Gehende und Fahrradfahrende sowie gleichermaßen viele Männer und Frauen den BÜ benutzten. Auf der anderen Seite legen die Ergebnisse der Studienarbeit von Illgen (2020) anhand der selben Stichprobe wie in der vorliegenden Arbeit jedoch auch nahe, dass es wahrscheinlich keine Unterschiede bei der Maßnahmenwirksamkeit bezüglich des Fortbewegungsmittels oder des Geschlechts gibt.

Ein weiterer Kritikpunkt ist bei der näheren Analyse der Übereinstimmungen zwischen den beiden Annotierenden ist aufgefallen. Es wurden die Fälle genauer untersucht, bei denen die Beurteilungen am stärksten voneinander abwichen: Bei zwölf beobachteten Personen urteilte ein Annotierender, dass kein Kontrollblick ausgeführt wurde und der andere Annotierende, dass sogar ein Kontrollblick in beide Richtungen gezeigt wurde. Zu Uneindeutigkeiten führten in fünf Fällen davon möglicherweise Accessoires im Gesichtsbereich wie eine Sonnenbrille, eine Brille mit breitem Rahmen oder Bartwuchs, da hier durch die geringe Auflösung der Videoaufnahmen Verpixelungen entstehen, wenn sich die Personen in der Ferne befinden. Zusätzliche Eigenbewegungen bei zwei Personen durch Joggen und durch Schlangenlinienfahren mit dem Fahrrad könnten ebenfalls dazu geführt haben, dass die Kopfausrichtung auf Grund der Videoqualität in der Ferne nicht eindeutig identifizierbar ist. In vier Fällen waren die Blickrichtung oder -ziel uneindeutig: Schaut die Versuchsperson zu dem Hund an der Straßenseite oder nach einem Schienenfahrzeug? Bei drei Personen könnte die fehlende Übereinstimmung der Annotierenden auf die Lichtverhältnisse bzw. auf die Videoqualität rückführbar sein, da das Gesicht durch Über- oder Unterbelichtung sowie durch einen starken Wechsel der Beleuchtung schwer erkennbar waren. Bei zukünftigen Studien sollte entweder noch intensiver darauf geachtet werden, dass die Kopfbewegung der beobachteten Personen eindeutig interpretierbar sind und keine Fehlinterpretationen erlauben oder ein Videoaufnahmeequipment ausgewählt werden, das eine bessere Aufnahmequalität aufweist. Außerdem zeigte sich bei drei Personen ein Kontrollblick genau in dem Bereich, in dem die Beobachtungszone begann oder endete. Je nach subjektiver Einschätzung des Annotierenden zählte dieser Blick oder nicht. Wäre in der Videodatei die Beobachtungszone beispielsweise farblich hervorgehoben, wären subjektive Fehleinschätzungen diesbezüglich weniger wahrscheinlich.

Eine weitere Limitation der Studienergebnisse ist, dass die Maßnahmen mit nur einer einwöchigen Pause implementiert und getestet wurden. Es ist nicht auszuschließen, dass positive Einflüsse auf das Kontrollblickverhalten der Querenden durch die Straßenmarkierung nachhaltig waren und somit die Vergleichsphase ohne Maßnahmen für die Wirksamkeitsprüfung des Blinklichts zu einer möglichen Überbewertung der Effektivität selbiger führte. Vor

der Implementierung des Blinklichts sollte nach einer längeren Erholungszeit eine zweite Kontrollerhebung durchgeführt werden. Dies hätte den Vorteil, dass mögliche Einschränkungen der internen Validität durch Zeitgescheheneffekte (z.B. ein gravierendes Zugunglück, über das in den Nachrichten berichtet wurde) oder durch exogenen Wandel (z.B. ein Einfluss durch Jahreszeiten) ebenfalls kontrolliert werden könnten. Eine andere Möglichkeit ist, an einem weiteren, vergleichbaren BÜ parallel beide Maßnahmen zu testen, wodurch entsprechende Einschränkungen der internen Validität ebenfalls entfallen. Ein Gewöhnungseffekt, der in der Regel bei derartigen Maßnahmen auftritt, konnte nicht gefunden werden. Dies ist nach nur zweiwöchiger Maßnahmenimplementierung nachvollziehbar, weshalb die Ergebnisse der vorliegenden Studie hinsichtlich des Ausbleibens nicht generalisierbar sind. Zwar ist es wahrscheinlich, dass ein Teil der zum späteren Messzeitpunkt beobachteten Personen den Maßnahmen bereits mehrfach zuvor begegnet waren, jedoch kann nicht ausgeschlossen werden, dass alle Beobachteten den Bahnübergang mit den Maßnahmen an diesem Tag zum ersten Mal überquerten. In diesem Zusammenhang wäre es relevant zu erfragen, wie regelmäßig die Querenden den BÜ nutzen, um nicht nur auf Grund der Uhrzeit bei der Beobachtung von regelmäßigen Nutzer:innen auszugehen. Eventuell ergeben sich durch Verbesserungen bei der Videotechnik zur Identifizierung von Personen Möglichkeiten, bei denen eine solche Erfassung automatisiert unter Beachtung geltender Datenschutzstandards stattfinden kann. Im Rahmen von Befragungen könnte überdies festgestellt werden, wie die Querenden die Maßnahmen subjektiv empfunden haben und ob die oben vorgeschlagenen Optimierungen an den Maßnahmen zur Salienz und zur Verständlichkeit sinnvoll sind.

Für Videoannotationsstudien empfiehlt Eby (2011), dass alle Annotierenden gemeinsam instruiert und trainiert werden. Im Rahmen einer konkreten Übung sollten die Ergebnisse unabhängig voneinander erfasst und anschließend gemeinsam diskutiert werden. Das gemeinsame Training sollte solange erfolgen, bis die Interrater-Reliabilität größer als 85 % ist. Grundsätzlich sind die gefundenen Interrater-Reliabilitäten auch ohne diese ausführliche Vorbereitung zufriedenstellend, was zeigt, dass das vorbereitete Trainingsmaterial mit den Anmerkungen zur Beobachtung und Protokollen offenbar geeignet, plausibel und obligat war.

5.6. Offene Forschungsfragen

Anknüpfungspunkte für zukünftige Forschungsfragen ergeben sich mit Blick auf die Ergründung von Langzeiteffekten auf die Maßnahmenwirksamkeit, die vorgeschlagenen Optimierungen der Maßnahmen, den Auswertungsumfang sowie den Erhebungsumfang.

Die längerfristige Erprobung der Maßnahmen zur Untersuchung möglicher Gewöhnungseffekte, wie sie bei einer dauerhaften Implementierung eintreten könnten, sollte bei zukünftigen Studien diskutiert werden. Besonders interessant wäre hierbei ein Vergleich der beiden Maßnahmen dahingehend, wann der erwartete Gewöhnungseffekt auftritt. Eventuell ergeben sich dadurch weitere Befunde bezüglich der Effektivität und Gestaltung der Maßnahmen bei der visuellen Aufmerksamkeitslenkung, da mit zunehmender Stimulusintensität die Habituation verzögert auftritt.

Im Rahmen der Maßnahmenoptimierung könnte die Straßenmarkierung salienter gestaltet werden, damit sie den Straßenverkehrsteilnehmenden stärker auffällt auch wenn sie außerhalb des Blickfelds aufgebracht wurde. Hinzu kommt, dass die Straßenmarkierung die Zielgruppe durch die Wahl eines Schriftzuges in deutscher Sprache einschränkt. Bei dem Blinklicht ist besonders kritisch anzumerken, dass es in dem derzeitigen Design zu Fehlinterpretationen kommen könnte: Das gelbe Blinken könnte so verstanden werden, dass sich tatsächlich ein Schienenfahrzeug nähert, was entsprechende maladaptive Lernprozesse begünstigt. Maßnahmenänderungen, die diese Kritikpunkte korrigieren, könnten in zukünftigen Studien praktisch erprobt und evaluiert werden.

In der vorliegenden Masterarbeit wurden nicht alle erhobenen Daten in vollem Umfang ausgewertet. Bei der Annotation wurden beispielsweise weitere Variablen erhoben (z.B. Alter der Personen), die bisher nicht ausgewertet wurden, womit zukünftige Forschungen differenziertere Aussagen zur Maßnahmeneffektivität bezüglich spezifischer Zielgruppen treffen können. Eine offene Forschungsfrage ist beispielsweise, ob die Altersgruppe der Querenden einen Einfluss auf die Maßnahmenwirksamkeit hat: Befindet sich in der Umgebung des BÜ ein Senior:innenheim könnten andere Maßnahmen als bei einem naheliegenden Schulzentrum bei der Mehrheit der Querenden zu einer höheren Wahrscheinlichkeit für ein adäquates Kontrollblickverhalten führen.

Eine maßgebliche Frage für den Rollout der entwickelten Maßnahmen ist außerdem, ob die getesteten Maßnahmen auch bei anderen Straßenverkehrsteilnehmenden wie Autofahrenden oder zweirädrig-motorisiert Fahrenden in der gleichen Weise wirken. In diesem Zusammenhang stellt sich außerdem die Frage, inwiefern die Platzierung im Raum und somit auch im Gesichtsfeld der den BÜ nutzenden Personen einen Einfluss auf die Wahrnehmung in verschiedenen Fortbewegungsmitteln hat. Diskutiert wurde bereits, ob mit der Zunahme der Geschwindigkeit in den Fahrzeugen die Erkennung der Straßenmarkierung abnimmt (Silla et al., 2019). Widerlegt werden konnte für die vorliegende Stichprobe durch die Auswertung von Illgen (2020) bereits, dass ein solcher Effekt für Fahrradfahrende im Vergleich mit den

langsameren zu Fuß Gehenden auftritt. Ein weiterer Aspekt blieb bislang unerforscht, der in Zusammenhang mit der Annahme steht, das Verhalten im Straßenverkehr sei regelbasiert, also teilweise automatisch ausgelöst: Welche Intervention ist bei welcher Gruppe bezogen auf die Querungsfrequenz besonders erfolgsversprechend? Die Frage „Wie oft überqueren Sie in der Regel den Bahnübergang?“ würde in Verbindung mit dem gezeigten Verhalten in der jeweiligen Bedingung Rückschlüsse darauf zulassen, ob bei den Personen, die den BÜ erstmals queren, die Entwicklung einer adäquaten Verhaltensstrategie durch die Maßnahmen gefördert werden können. Genauso wäre es interessant zu eruieren, wie die alltäglich Querenden, die aus Gewohnheit nicht adäquat schauen, dazu bewegt werden könnten, den Kontrollblick wieder auszuführen. Darüber hinaus wäre es beachtenswert, aus welchen Gründen ein gewisser Anteil der beobachteten Personen trotz der innovativen Maßnahmen kein adäquates Kontrollblickverhalten zeigt. Die vorliegende Masterarbeit liefert keine Erklärungen dafür, an welcher Stelle die Maßnahmen in den Testbedingungen im Rahmen kognitiver Informationsmodelle greifen und den adäquaten Kontrollblick anregen, was durch zukünftige Forschungen aufgegriffen werden könnte.

Wie eine Befragung z.B. zur Querungsfrequenz im Kontext des Ablaufs der vorliegenden Studie eingebaut werden könnte, ohne das natürliche Geschehen am BÜ zu beeinflussen und die Vorteile der verdeckten, nicht-teilnehmenden Feldbeobachtung mithilfe der Videoannotation zu schmälern, sollte sorgfältig geplant werden. Spannend wäre es außerdem, die Qualität des Kontrollblickverhaltens zu quantifizieren und zusätzliche Variablen wie Dauer, Häufigkeit oder Ausprägtheit der Kopfbewegung mit zu erfassen. Die Idee hinter dieser Erweiterung der Kontrollblickerfassung ist, dass mit der Zunahme der Maßnahmeneffektivität auch die Schutzmotivation steigt, was zu entsprechend ausgeprägterem Sicherheitsverhalten führt. Denkbar wäre in dem Zusammenhang auch, dass es einen Zusammenhang dieser Qualitätskriterien mit der Eindeutigkeit bei der Beurteilung von Außenstehenden gibt, was wiederum Auswirkungen auf die Interrater-Reliabilität hat.

Insgesamt konnte die vorliegende Masterarbeit offene Fragen der bereits durchgeführten Auswertung von Illgen (2020) zur Güte der gewählten Datenerhebungs- und Auswertungsmethode klären. Die Ergebnisse und Überlegungen zu deren Interpretation sowie darüber hinaus neue Ideen zur Optimierung der entwickelten Maßnahmen lassen Spielraum für zukünftige Studien und Forschungsarbeiten zu dem spannenden Forschungsfeld, wie mit Hilfe innovativer und kostengünstiger Maßnahmen die Sicherheit an Bahnübergängen verbessert werden kann. Die zielgruppengerechte Anpassung der Umwelt und technischer Systeme zur Förderung wünschenswerter Verhaltensweisen und zur Vermeidung von Fehlern sind die

maßgeblichen Ziele der Human-Factors-Forschung, die eine wichtige Komponente in dem Projekt SAFER-LC war. Das im Projekt erarbeitete Portfolio mit Empfehlungen zur Gestaltung von BÜ zur Erhöhung der Sicherheit im Kreuzungsbereich zwischen Straßen- und Schienenverkehr leistet somit einen wertvollen Beitrag für die Vision Zero: Die Vermeidung von Toten und Verletzten durch eine sichere Gestaltung der Verkehrsinfrastruktur.

6. Literaturverzeichnis

- Abdi, H. (2010). Holm's Sequential Bonferroni Procedure (Neil Salkind, Hrsg.) (*Encyclopedia of Research Design*). Thousand Oaks, CA: Sage. Verfügbar unter <https://personal.utdallas.edu/~herve/abdi-Holm2010-pretty.pdf>
- Abrahamse, W. & Matthies, E. Informational strategies to promote pro-environmental behaviours: Changing knowledge, awareness, and attitudes. In *Environmental psychology: An Introduction* (S. 223-232).
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50 (2), 179-211.
- Bortz, J. & Schuster, C. (2010). *Statistik. Für Human- und Sozialwissenschaftler: mit 163 Tabellen* (7. Aufl.). Berlin: Springer.
- Bundesministerium der Justiz und für Verbraucherschutz. (2020). Gesetze im Netz, Bundesministerium der Justiz und für Verbraucherschutz. Zugriff am 26.10.2020. Verfügbar unter <http://www.gesetze-im-internet.de/impressum.html>
- Cian, L., Krishna, A. & Elder, R. S. (2015). A Sign of Things to Come: Behavioral Change through Dynamic Iconography. *Journal of Consumer Research*, 41 (6), 1426-1446.
- Conzola, V. C. & Wogalter, M. S. (2001). A Communication-Human Information Processing (C-HIP) approach to warning effectiveness in the workplace. *Journal of Risk Research*, 4 (4), 309-322.
- Cox, B. S., Cox, A. B. & Cox, D. J. (2000). Motivating signage prompts safety belt use among drivers exiting senior communities. *Journal of applied behavior analysis*, 33 (4), 635-638.
- DB Netze. (2018). *Bahnübergänge im Spiegel der Statistik - Bahnübergangsstatisik 2018: Version für Dritte* (DB Netz AG / TÜV Süd Industrie Service GmbH, Hrsg.), Berlin, München.
- Demuth, C. (2020). Videoanalysen. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren* (S. 751-771). Wiesbaden: Springer Fachmedien Wiesbaden.
- Deutsche Bahn AG. (2019). Schranken schließen nicht ohne Grund! Unfälle durch richtiges Verhalten vermeiden! Zugriff am 10.09.2020. Verfügbar unter https://www.deutschebahn.com/resource/blob/1173482/8f56ffa5fe8c3824e4fde735e9fcb26c/unfaelle_bahnueb-data.pdf
- Dressler, A. & Gripenkoven, J. (2018). SAFER-LC: Innovative Lösungen für mehr Sicherheit am Bahnübergang. *EI - Der Eisenbahningenieur*.
- Dressler, A., Illgen, J. & Gripenkoven, J. (2020). Is a train coming? Effects of two low-cost safety measures on road user behavior at a passive level crossing. German RHF 2020 - 4th German Workshop on Rail Human Factors. 2. Dez. 2020, Virtuelle Konferenz.
- Dressler, A., Silla, A., Kortsari, A., Havarneanu, G., Whalley, S. & Gripenkoven, J. (2020, April). Human-centered measures to enhance safety at level crossings. Proceedings of 8th Transport Research Arena TRA 2020, Helsinki, Finland.
- Eby, D. W. (2011). Chapter 5 - Naturalistic Observational Field Techniques for Traffic Psychology Research. In B. E. Porter (Hrsg.), *Handbook of Traffic Psychology* (S. 61-72). San Diego: Academic Press. Verfügbar unter <http://www.sciencedirect.com/science/article/pii/B9780123819840100050>

- Eisenbahn-Bundesamt. (2018). Bericht des Eisenbahn-Bundesamts. gemäß Artikel 18 der Richtlinie über Eisenbahnsicherheit in der Gemeinschaft (Richtlinie 2004/49/EG, „Sicherheitsrichtlinie“) über die Tätigkeiten als Sicherheitsbehörde.
- Esser, H. (1990). „Habits“, „Frames“ und „Rational Choice“. *Zeitschrift für Soziologie*, 19 (4).
- European Union Agency for Railways (Publications Office of the European Union, Hrsg.). (2018). Railway Safety in the European Union. Safety overview 2017. Verfügbar unter https://www.era.europa.eu/sites/default/files/library/docs/safety_interoperability_progress_reports/railway_safety_performance_2017_en.pdf
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39 (2), 175-191.
- Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics (5. edition)*. London: SAGE.
- Glendon, A. I. & Walker, B. L. (2013). Can anti-speeding messages based on protection motivation theory influence reported speeding intentions? *Accident; analysis and prevention*, 57, 67-79.
- Google Maps (Google, Hrsg.). (2020). Verfügbar unter <http://maps.google.com/maps>
- Grippenkoven, J. & Dietsch, S. (2015). Gaze direction and driving behavior of drivers at level crossings. *Journal of Transportation Safety & Security*, 8, 0.
- Grippenkoven, J., Thomas, B. & Lemmer, K. (2016). PeriLight - effektive Blicklenkung am Bahnübergang. *Eisenbahningenieur*, 48-51.
- Groeger, J. A. (2000). *Understanding driving. Applying cognitive psychology to a complex everyday task (Frontiers of cognitive science)*. Hove: Psychology Press.
- Groves, P. M. & Thompson, R. F. (1970). Habituation: a dual-process theory. *Psychological Review*, 77(5), 419–450.
- Hussy, W., Echterhoff, G. & Schreier, M. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin, Heidelberg: Springer.
- IBM SPSS Statistics for Windows (2011) [Computer software]. Armonk, New York: IBM Corporation.
- Illgen, J. (2020). Wirkung kostengünstiger Sicherungsmaßnahmen auf das Verhalten gefährdeter Straßennutzer an Bahnübergängen. Unveröffentlichte Studienarbeit. Braunschweig: TU Braunschweig.
- Klößner, C. A. (2015). *The Psychology of Pro-Environmental Communication*. London: Palgrave Macmillan UK.
- Knake-Langhorst, S., Gimm, K., Frankiewicz, T. & Köster, F. (2016). Test Site AIM – Toolbox and Enabler for Applied Research and Development in Traffic and Mobility. *Transportation Research Procedia*, 14, 2197-2206.
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33 (1), 159.
- Lippke, S. & Renneberg, B. (2011). Theorien und Modelle des Gesundheitsverhaltens. In B. Renneberg (Hrsg.), *Gesundheitspsychologie*. 35-60. Heidelberg: Springer.

- Matthies, E., Kastner, I., Klesse, A. & Wagner, H.-J. (2011). High reduction potentials for energy user behavior in public buildings: how much can psychology-based interventions achieve? *Journal of Environmental Studies and Sciences*, 1 (3), 241-255.
- Max Planck. (2020) ELAN [Computer software]. Nijmegen.
- Microsoft Office Home and Business 2016 (2016) [Computer software]: Microsoft.
- R. A language and environment for statistical computing: reference index. (2010). Vienna: R Foundation for Statistical Computing.
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13 (3), 257-266.
- Reason, J. T. (1990). *Human error*. Cambridge, England: Cambridge University Press.
- Richter, S. (Technischen Universität Dresden, Hrsg.). (2016). *Verkehrspsychologie – Verkehrspädagogik. Eine Einführung für Lehramtsstudierende*. Verfügbar unter https://tu-dresden.de/bu/verkehr/ivs/vpsy/ressourcen/dateien/publikationen/Lehrhandbuch_Stiftung_GE_10-16_FINAL_277331.pdf?lang=de
- Rogers, R., Cacioppo, J. & Petty, R. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In *Social Psychophysiology*, 153-177, New York: Guilford Press.
- Rudin-Brown, C. M., George, M. F.-S. & Stuart, J. J. (2014). Human Factors Issues of Accidents at Passively Controlled Rural Level Crossings. *Transportation Research Record: Journal of the Transportation Research Board*, 2458 (1), 96-103.
- Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84 (1), 1-66.
- Schumacher, O. & Weihgold, M. (Deutsche Bahn AG, Hrsg.). (2017). Bahnübergänge. Wussten Sie schon, dass... Verfügbar unter https://www.deutschebahn.com/resource/blob/1173476/9266e88e8e740a2af29431cbf7282784/wussten_sie_bahnuebergang-data.pdf
- Schwarz, N. (1985). Theorien konzeptgesteuerter Informationsverarbeitung in der Sozialpsychologie. *Theorien der Sozialpsychologie*, 3, 269-291. Verfügbar unter https://homepage.univie.ac.at/michael.trimmel/kogpsych_ss2002/Theorien_konzeptgesteuerter_Informationsverarbeitung.pdf
- Silla, A., Virtanen, A., Lehtonen, E., Boufidis, N., Salanova Grau, J., Dressler, A. et al. (2019). Results of the evaluation of the pilot tests: SAFER-LC project Deliverable D4.4.
- Städtler, T. (1998). *Lexikon der Psychologie. Wörterbuch, Handbuch, Studienbuch*. Stuttgart, Deutschland: Alfred Körner Verlag.
- Steg, L. (2013). *Environmental psychology. An introduction (BPS textbooks in psychology)*. Chichester: Wiley-Blackwell.
- Thompson, R. F. (2009). Habituation: A history. *Neurobiology of Learning and Memory*, 92(2), 127–134.
- Wirtz, M. & Kutschmann, M. (2007). Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Maße. *Die Rehabilitation*, 46 (6), 1-8.
- Wogalter, M. S. & Laughery, K. R. (1996). WARNING! Sign and Label Effectiveness. *Current Directions in Psychological Science*, 5 (2), 33-37.

-
- Wolfe, J. M. & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews. Neuroscience*, 5(6), 495–501.
- Zachrisson, J. & Boks, C. (2010). When to apply different design for sustainable behaviour strategies?, Proceedings of Knowledge Collaboration & Learning for Sustainable Innovation ERSCP-EMSU conference, Delft, Niederlande, October 25-29, 2010.

Anhang A

Häufigkeitstabelle des differenzierten Kontrollblickverhaltens

Tabelle 4

Häufigkeiten des differenzierten Kontrollblickverhaltens innerhalb der Bedingungen

Testbedingung	differenzierter Kontrollblick [% innerhalb der Testbedingung]			N _{Testbedingung}
	gar nicht	in eine Richtung	beide Richtungen	
Baseline	39.6	22.9	37.5	48
Markierung	15.6	22.9	61.5	96
Blinklicht	12.5	12.5	75.0	96

Anhang B

Pearson Chi-Quadrat-Test zur Maßnahmeneffektivität

Tabelle 5

Häufigkeitstabelle der Analysen zur Maßnahmeneffektivität

		Kontrollblick		Total	
		inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen		
Markierung vs. Baseline	Markierung	Anzahl	37	59	96
		Erwartete Anzahl	44.7	51.3	96
		% innerhalb der Bedingung	38.5	61.5	100
		Stand. Residuum	-1.1	1.1	
	Baseline	Anzahl	30	18	48
		Erwartete Anzahl	22.3	25.7	48
		% innerhalb der Bedingung	62.5	37.5	100
		Stand. Residuum	1.6	-1.5	
Blinklicht vs. Baseline	Blinklicht	Anzahl	24	72	96
		Erwartete Anzahl	36	60	96
		% innerhalb der Bedingung	25.0	75.0	100
		Stand. Residuum	-2.0	1.5	
	Baseline	Anzahl	30	18	48
		Erwartete Anzahl	18	30	48
		% innerhalb der Bedingung	62.5	37.5	100
		Stand. Residuum	2.8	-2.2	
Markierung vs. Blinklicht	Markierung	Anzahl	37	59	96
		Erwartete Anzahl	30.5	65.5	96
		% innerhalb der Bedingung	38.5	61.5	100
		Stand. Residuum	1.2	-0.8	
	Blinklicht	Anzahl	24	72	96
		Erwartete Anzahl	30.5	65.5	96
		% innerhalb der Bedingung	25.0	75.0	100
		Stand. Residuum	-1.2	0.8	

Tabelle 6

Ergebnisse der Chi-Quadrat-Tests zur Maßnahmeneffektivität

		Testwert	df	p (einseitiger Test) ^a
Markierung vs. Baseline	Pearson Chi-Quadrat	7.383	1	.008**
	Cramérs-V	.226		
	Anzahl der gültigen Fälle	144		
Blinklicht vs. Baseline	Pearson Chi-Quadrat	19.2	1	< .001***
	Cramérs-V	.365		
	Anzahl der gültigen Fälle	144		
Markierung vs. Blinklicht	Pearson Chi-Quadrat	4.061	1	.022*
	Cramérs-V	.145		
	Anzahl der gültigen Fälle	192		

Anmerkung. * p < .05 = signifikant; ** p < .01 hoch signifikant; *** p < .001 höchst signifikant.

^a nach Bonferroni-Holm korrigiert

Anhang C

Häufigkeitstabelle zum adäquaten Kontrollblickverhalten

Tabelle 7

Häufigkeitstabelle zum adäquaten Kontrollblickverhalten der acht Messtage der Testbedingungen mit Maßnahmen

			Kontrollblick [% innerhalb des Messtages]		N _{Messtag}
			inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen	
Testbedingung und Messtage mit entsprechendem Wochentag	Markierung	3. Di	41.7	58.3	24
		4. Do	37.5	62.5	24
		5. Di	37.5	62.5	24
		6. Do	37.5	62.5	24
	Blinklicht	7. Di	20.8	79.2	24
		8. Do	25.0	75.0	24
		9. Di	29.2	70.8	24
		10. Do	25.0	75.0	24

Anhang D

Pearson Chi-Quadrat-Test zum Gewöhnungseffekt

Tabelle 8

Häufigkeitstabellen der Analysen zum Gewöhnungseffekt der beiden Testphasen mit Maßnahmen

			Kontrollblick		Total
			inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen	
Markierung	erster Messtag	Anzahl	10	14	24
		Erwartete Anzahl	9.5	14.5	24
		% innerhalb des Messtages	41.7	58.3	100
		Stand. Residuum	0.2	-0.1	
	letzter Messtag	Anzahl	9	15	24
		Erwartete Anzahl	9.5	14.5	24
		% innerhalb des Messtages	37.5	62.5	100
		Stand. Residuum	-0.2	0.1	
Blinklicht	erster Messtag	Anzahl	5	19	24
		Erwartete Anzahl	5.5	18.5	24
		% innerhalb des Messtages	20.8	79.2	100
		Stand. Residuum	-0.2	0.1	
	letzter Messtag	Anzahl	6	18	24
		Erwartete Anzahl	5.5	18.5	24
		% innerhalb des Messtages	25.0	75.0	100
		Stand. Residuum	0.2	-0.1	

Tabelle 9

Ergebnisse der Chi-Quadrat-Tests zu den Gewöhnungseffekten

Testbedingung		Testwert	df	p (einseitiger Test) ^a
Markierung (erster vs. letzter Messtag)	Pearson Chi-Quadrat	0.087	1	1.
	Anzahl der gültigen Fälle	48		
Blinklicht (erster vs. letzter Messtag)	Pearson Chi-Quadrat	0.118	1	1.
	Anzahl der gültigen Fälle	48		

Anmerkung.

^a nach Bonferroni-Holm korrigiert

Anhang E

Pearson Chi-Quadrat-Test zum Wochentageffekt

Tabelle 10

Häufigkeitstabellen der Analysen zum Wochentageffekt der beiden Testphasen mit Maßnahmen

		Kontrollblick		Total	
		inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen		
Markierung	Dienstag	Anzahl	19	29	48
		Erwartete Anzahl	18.5	29.5	48
		% innerhalb des Wochentages	39.6	60.4	100
		Stand. Residuum	0.1	-0.1	
	Donnerstag	Anzahl	18	30	48
		Erwartete Anzahl	18.5	29.5	48
		% innerhalb des Wochentages	37.5	62.5	100
		Stand. Residuum	-0.1	0.1	
Blinklicht	Dienstag	Anzahl	12	36	48
		Erwartete Anzahl	12	36	48
		% innerhalb des Wochentages	25.0	75.0	100
		Stand. Residuum	0.0	0.0	
	Donnerstag	Anzahl	12	36	48
		Erwartete Anzahl	12	36	48
		% innerhalb des Wochentages	25.0	75.0	100
		Stand. Residuum	0.0	0.0	

Tabelle 11

Ergebnisse der Chi-Quadrat-Tests zu den Wochentageffekten

Testbedingung	Testwert	df	p (zweiseitiger Test) ^a	
Markierung (Dienstage vs. Donnerstage)	Pearson Chi-Quadrat	0.044	1	1.
	Anzahl der gültigen Fälle	96		
Blinklicht (Dienstage vs. Donnerstage)	Pearson Chi-Quadrat	0.000	1	1.
	Anzahl der gültigen Fälle	96		

Anmerkung.

^a nach Bonferroni-Holm korrigiert

Anhang F

Interrater-Reliabilitäten zu den Eigenschaften der VRU

Tabelle 12

Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich des Geschlechts der VRU

			Annotation 2: Geschlecht		Total
			männlich	weiblich	
Annotation 1: Geschlecht	männlich	Anzahl	115	5	120
		% von Total	48.1	2.1	50.2
	weiblich	Anzahl	5	114	119
		% von Total	2.1	47.7	49.8
Total	Anzahl	120	119	239	
	% von Total	50.2	49.8	100.0	

Tabelle 13

Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich der Altersgruppe der VRU

			Annotation 2: Altersgruppe			Total
			Jugendliche	Erwachsene	Senior:innen	
Annotation 1: Altersgruppe	Jugendliche	Anzahl	10	3	0	13
		% von Total	4.2	1.2	0.0	5.4
	Erwachsene	Anzahl	16	172	19	207
		% von Total	6.7	71.7	7.9	86.2
	Senior:innen	Anzahl	0	2	18	20
		% von Total	0.0	0.8	7.5	8.3
Total	Anzahl	26	177	37	240	
	% von Total	10.8	73.8	15.4	100	

Tabelle 14

Häufigkeitstabelle zur Übereinstimmung zwischen Annotation 1 und 2 bezüglich der Fortbewegungsart der VRU

		Annotation 2: VRU-Typ			Total	
		zu Fuß gehend	radfahrend	andere		
Annotation 1: VRU-Typ	zu Fuß gehend	Anzahl	115	2	0	117
		% von Total	47.9	0.8	0.0	48.8
	radfahrend	Anzahl	4	114	1	119
		% von Total	1.7	47.5	0.4	49.6
	andere	Anzahl	0	0	4	4
		% von Total	0.0	0.0	1.7	1.7
Total	Anzahl	26	119	116	240	
	% von Total	10.8	49.6	48.3	100	

Anmerkung. VRU-Typ = vulnerable road user - Typ (entspricht dem Fortbewegungsmittel der Versuchsperson).

Anhang G

Interrater-Reliabilitäten zum Kontrollblickverhalten der VRU

Tabelle 15

Häufigkeitstabelle zur Übereinstimmung der Beurteilung adäquaten Kontrollblickverhaltens zwischen Annotation 1 und 2

		Annotation 2: adäquater Kontrollblick			Total
		inadäquat: gar nicht oder in eine Richtung	adäquat: beide Richtungen		
Annotation 1: adäquater Kontrollblick	inadäquat: gar nicht oder in eine Richtung	Anzahl	73	28	101
		% von Total	30.4	11.7	42.1
	adäquat: in beide Richtungen	Anzahl	18	121	139
		% von Total	7.5	50.4	57.9
Total	Anzahl	91	149	239	
	% von Total	37.9	62.1	100	

Anmerkung. Prozentuale Angaben zur Übereinstimmung sind zur Verbesserung der Übersichtlichkeit fett hervorgehoben und Zellen mit übereinstimmenden Urteilen der beiden Annotierenden grau hinterlegt.

Tabelle 16

Häufigkeitstabelle zur Übereinstimmung der Beurteilung des differenzierten Kontrollblickverhaltens zwischen Annotation 1 und 2

		Annotation 2: differenzierter Kontrollblick			Total	
		gar nicht	in eine Richtung	beide Richtungen		
Annotation 1: differenzierter Kontrollblick	gar nicht	Anzahl	37	19	5	61
		% von Total	15.4	7.9	2.1	25.4
	in eine Richtung	Anzahl	2	15	23	40
		% von Total	0.8	6.2	9.6	16.7
	in beide Richtungen	Anzahl	7	11	121	139
		% von Total	2.9	4.6	50.4	57.9
Total	Anzahl	46	45	149	240	
	% von Total	19.2	18.8	62.1	100	

Anmerkung. Prozentuale Angaben zur Übereinstimmung sind zur Verbesserung der Übersichtlichkeit fett hervorgehoben und Zellen mit übereinstimmenden Urteilen der beiden Annotierenden grau hinterlegt.

Eidesstattliche Erklärung

Hiermit erkläre ich, Jennifer Nöske, geb. Lackner am 28. Mai 1990 in Lübeck, dass ich die vorliegende schriftliche Masterarbeit „'Bahn frei?': Effektive Förderung des Kontrollblicks an Bahnübergängen und Analyse der Objektivität der gewählten Untersuchungsmethode“ selbst angefertigt und keine anderen als die angegebenen Hilfsmittel genutzt habe.

Alle Stellen, die wortwörtlich oder dem Sinn nach aus anderen Publikationen entnommen sind, habe ich unter exakter Angabe des Quelltextes eindeutig als Entlehnung gekennzeichnet.

Die Arbeit ist in gleicher oder ähnlicher Form noch nicht als Prüfungsarbeit eingereicht worden.

Ort, Datum

Unterschrift