

## **Die Performance einer Webapplikation als User Experience Problem**

Verbesserungsmöglichkeiten und Lösungen

Autoren:

**Cristina Menéndez Carmen Rickenbach Tobias Schmidt**

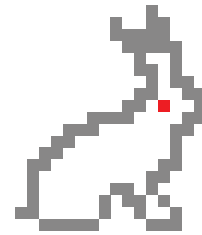


Alice:

# How long is forever?

White Rabbit:

Sometimes, just one second.



# Impressum

Januar 2017

Im Rahmen des Masters of Advanced Studies in Human Computer Interaction Design an der HSR Hochschule für Technik Rapperswil und der Universität Basel.

Projektteam:

Cristina Menéndez | Carmen Rickenbach | Tobias Schmidt

Auftraggeber:

SBB AG

Eva Siegenthaler

Betreuer:

Marc Blume

In der vorliegenden Arbeit wird der Einfachheit halber nur die männliche Form verwendet. Die weibliche Form ist selbstverständlich immer miteingeschlossen.

# Danksagung

Unser grosser Dank gebührt SBB, insbesondere Eva Siegenthaler und Mike Morgenthaler für die Erteilung des Auftrages und die Unterstützung während der Masterarbeit. Auch vielen Dank an Andreas Elsener, Zürcher Kantonalbank, für die ursprüngliche Idee, die zu dieser Masterarbeit geführt hat.

Wir bedanken uns herzlich bei unserem Betreuer Marc Blume. Er hat uns kompetent unterstützt, vor allem bei den wissenschaftlichen Aspekten unserer Arbeit war er ein wichtiger Berater.

Ein besonderer Dank geht an Svenja, Iris, Marco, Andreas, Manuel, Carina, Barbara, Claudia, Sybille, Laura, Hans, Georgia und Karin für die Hilfe mit Kontakten, Reviews, Tipps und Statistiken. Auch vielen Dank an die 282 Personen, die an der Umfrage teilgenommen haben.

Zum Schluss wollen wir uns bei unserer Familie, Partnern, Freunden und Arbeitgebern für das Verständnis und die Unterstützung der letzten neun Monaten ganz herzlich bedanken.

# Abstract

Die Arbeit beschäftigt sich mit Massnahmen zur Verbesserung der User Experience des Ladevorgangs bei Webapplikation. Insbesondere werden Ladeanzeigen untersucht, mit deren Hilfe zum einen die durch die Benutzer wahrgenommene Ladezeit verkürzt und zum anderen die allgemeine User Experience und Benutzerzufriedenheit optimiert werden soll. Als konkreter Anwendungsfall dient dabei der Prototyp des neuen SBB-Ticketshops als Webapplikation. Aufgrund theoretischer Grundlagen wird hier das Zusammenspiel der Faktoren Zeit, Typ der Ladeanzeige, Position sowie Ästhetik erörtert und untersucht. Am konkreten Fall werden hypothetische Verbesserungen erarbeitet, die dann in einer experimentellen Studie überprüft werden.

Für den konkreten Fall finden Benutzer am Prototyp ohne Fortschrittsanzeige keinen Gefallen und die Wartezeit wird von den Benutzern als unangemessen eingestuft. Dieses Resultat deckt sich mit der Literatur. Im Vergleich zwischen vier Alternativen Prototypen sind Verständlichkeit und Voraussagbarkeit beim Prototypen mit Fortschrittsbalken am besten und beim Prototypen ohne Feedback am schlechtesten. Die visuelle Ästhetik (sympathisch, originell, spannend, attraktiv) ist bei spielerischer Fortschrittsanzeige am besten. In diesem Kontext konnte kein Zusammenhang zwischen der empfundene Ladedauer und der Art einer Fortschrittsanzeige nachgewiesen werden.

Die Ergebnisse und Empfehlungen sollen sowohl zur Verbesserung des SBB-Ticketshops dienen als auch der UX-Gemeinschaft allgemeine Erkenntnisse zum Design und Verwendung von Ladeanzeigen liefern.

# Selbständigkeitserklärung

Hiermit bestätigen wir,

- dass wir die vorliegende Arbeit selber und ohne fremde Hilfe durchgeführt haben, mit Ausnahme:
  - der Bestimmung der statistischen Testverfahren
  - der Berechnung der Statistiken und der G-Power der Stichprobe in den Anhängen 1, 2 und 3
- dass wir sämtliche verwendeten Quellen erwähnt und gemäss gängigen wissenschaftlichen Regeln korrekt zitiert haben, und
- dass wir keine durch Copyright geschützten Materialien (z.B. Bilder) in dieser Arbeit in unerlaubte Weise genutzt haben,
- dass wir in dieser Arbeit keine Adressen, Telefonnummern und andere persönliche Daten von Personen, die nicht zum Kernteam gehören, publizieren.

Olten, 30. Januar 2017

**Cristina Menéndez**

Zürich, 30. Januar 2017

**Carmen Rickenbach**

Konstanz, 30. Januar 2017

**Tobias Schmidt**

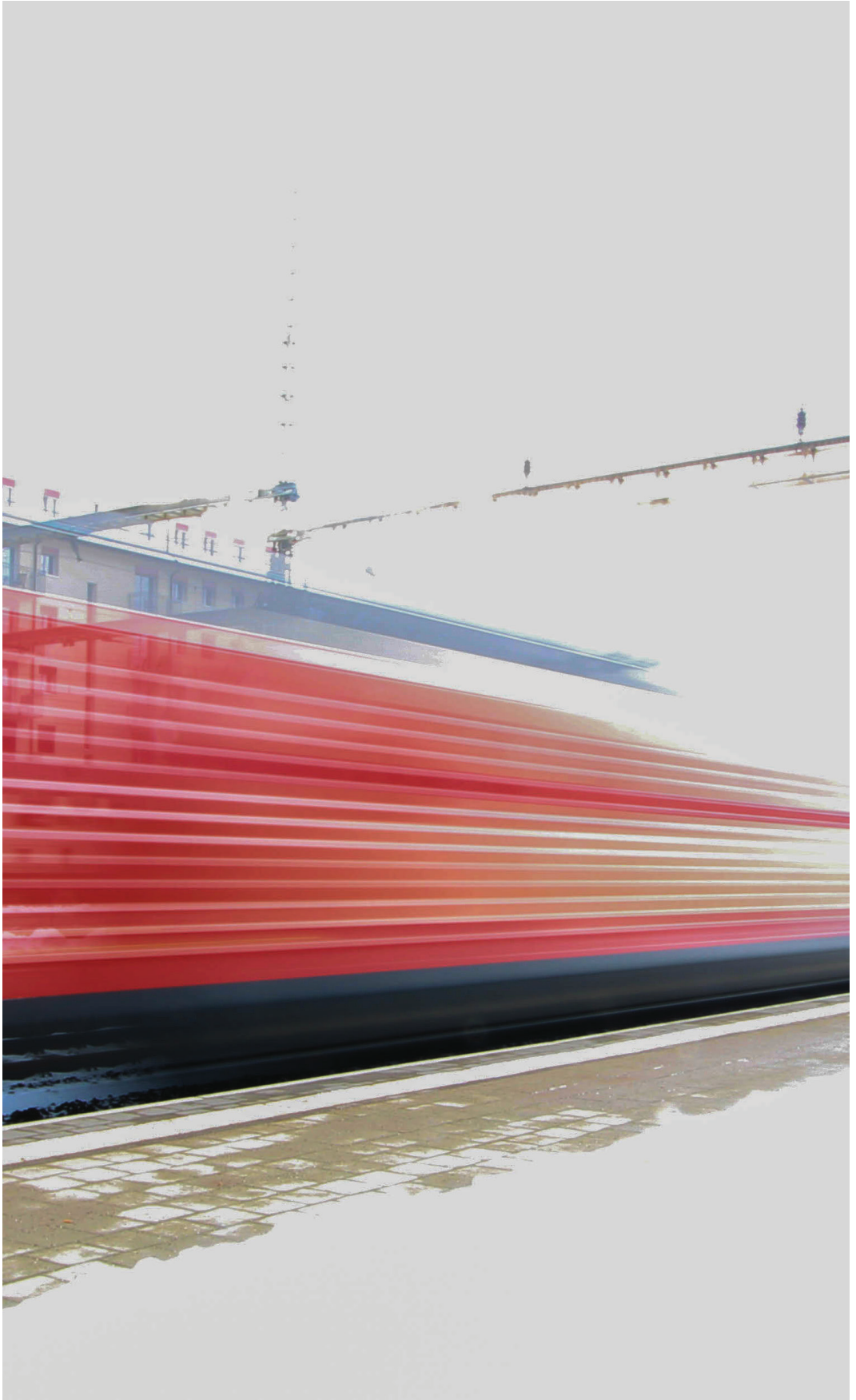
# Inhaltverzeichnis

<b>1. Einleitung</b>	<b>13</b>
1.1 Auftraggeber und Ausgangslage	13
1.2 Ziel der Arbeit	13
1.3 Abgrenzung	14
1.4 Allgemeines Vorgehen	14
<b>2. Theoretischer Hintergrund</b>	<b>17</b>
2.1 Performance	17
2.1.1 Das Ladeverhalten im Kontext der ISO 9241-110 Richtlinien	17
2.1.2 Wahrgenommene Performance	18
2.2 Responsiveness	19
2.2.1 Fortschrittsanzeigen	21
2.3 Visuelle Ästhetik des Ladeverhaltens	23
2.4 Bedeutung des theoretischen Hintergrunds für unsere Arbeit	24
<b>3. Vorbereitung der Studie</b>	<b>27</b>
3.1 User Centered Design-Prinzipien	27
3.2. Analyse und Spezifikation des Kontexts des Systems	27
3.2.1 Stakeholder Interview	28
3.2.2 Konkurrenzanalyse	31
<b>4. Stadien der wissenschaftlichen Methode</b>	<b>35</b>
4.1 Hypothesenbildung	36
4.2 Statistischer Versuchsplan	36
4.3 Versuchsaufbau	38

4.4 Versuchsdurchführung	40
4.5. Auswertung (Datenanalyse)	41
4.6 Schlussfolgerung (Antworten)	41
<b>5. Versuchsaufbau</b>	<b>43</b>
5.1 Definition von User Experience in unserer Arbeit	43
5.2 Designvarianten	44
5.2.1 Zeit	44
5.2.2 Positionierung der Fortschrittsanzeige	44
5.2.3 Typ der Fortschrittsanzeige	46
5.2.4. Unsere Variable für die Studie	47
5.3 Hypothesenformulierung	48
<b>6. Durchführung der Studie</b>	<b>53</b>
6.1 Prototyp	53
6.1.1 Ausarbeitungsgrad	53
6.1.2 Technische Anforderungen	53
6.1.3 Benutzerführung im Remote-Testing	54
6.2 Vorversuch: Usability-Test	55
6.2.1 Vorbereitung: Cognitive Walkthrough	57
6.2.2 Durchführung: Usability-Test	59
6.2.3 Nachbearbeitung: Erkenntnisse aus den Usability-Tests und dem Hallway Testing	63
6.3 Hauptversuch: Online-Umfrage	65
6.3.1 Fragebogen	65
6.3.2 Durchführung der Online-Umfrage	68

<b>7. Ergebnisse</b>	<b>73</b>
7.1 Überarbeitung des Datensatzes	73
7.2 Deskriptive Daten	74
7.3 Validierung der Hypothesen	78
7.3.1 Der Einfluss der Fortschrittsanzeige auf die empfundene Ladedauer hat sich nicht bestätigt	78
7.3.2 Die visuelle Ästhetik ist bei spielerischer Fortschrittsanzeige am besten	79
7.3.3 Schlechteste Benutzerzufriedenheit bei Ladedauer ohne Fortschrittsanzeige	81
7.3.4 Die Aussagekraft ist für Prototyp 1 am schwächsten, die Verständlichkeit und Voraussagbarkeit ist bei Prototyp 2 am stärksten	82
<b>8. Diskussion</b>	<b>85</b>
8.1 Empfehlung an die SBB	86
8.2 Exkurs: Expert-Review Ergebnisse	87

<b>9. Reflexion</b>	<b>89</b>
<b>10. Literaturverzeichnis</b>	<b>90</b>
<b>11. Abbildungsverzeichnis</b>	<b>96</b>
<b>12. Glossar</b>	<b>98</b>
<b>13. Anhang</b>	<b>103</b>



# 1. Einleitung

Diese Masterarbeit wird im Rahmen des Studiums zum Master of Advanced Studies in Human Computer Interaction Design der HSR Hochschule für Technik Rapperswil und der Universität Basel erstellt.

Auf der Suche nach einem Thema für die Masterarbeit trafen wir mit den SBB in Kontakt. Die SBB zeigten Interesse und boten uns verschiedene Themen an. Das Thema *Performance und Wechselwirkung mit UX* weckte unser Interesse am meisten und auch die SBB signalisierten, dass ihnen dieses Thema im Rahmen eines laufenden Projekts auch einen konkreten Nutzen generieren könnte. So entschieden wir uns, in dieses Thema einzusteigen, obwohl wir wussten, dass vieles Neuland sein würde. Gerade darin sahen wir aber den grossen Reiz, uns der Herausforderung zu stellen.

## 1.1

### Auftraggeber und Ausgangslage

Die Auftraggeberin der Arbeit sind die SBB, vertreten durch die Abteilung UX. Unsere Ansprechpartnerin und Begleiterin ist Eva Siegenthaler, Leiterin der Abteilung UX.

Der Ticketkauf bei den SBB, sei es am Billettautomaten oder online, gilt als eines der bekanntesten Usability Probleme der Schweiz (Mair-Noack, 2007). Die Komplexität und die Probleme in der Bedienung wurden sogar in den Medien diskutiert. Die SBB erkannten das Problem und starteten ein Projekt für das Redesign des Webshops, um so auch die wichtigsten Usability-Probleme der vorherigen Version zu adressieren. Nicht nur das Frontend des Webshops (Abbildung 1.1) wurde komplett überarbeitet, sondern auch externe und interne Backend-Systeme überdacht und überarbeitet.

Im Laufe dieser Entwicklung wurden seitens der SBB Probleme mit den Ladezeiten festgestellt. Das Zusammenspiel verschiedener Backends unterschiedlicher Anbieter macht die Webapplikation nicht nur komplex, sondern auch anfällig für Performance-Probleme, für die es keine unmittelbare technische Lösung gibt.

Hier kommt unsere Masterarbeit zum Zug: Wir gehen davon aus, dass eine schlechte Performance die User Experience (UX) potenziell negativ beeinflusst. Was muss geschehen, damit der Benutzer auch bei schlechter Performance die Experience mehrheitlich trotzdem positiv wertet? Mit welchen UX-Massnahmen kann die Experience verbessert werden?

## 1.2

### Ziel der Arbeit

Mit dieser Arbeit sollen

- geeignete UX-Massnahmen identifiziert werden, die die User Experience bei hohen Ladezeiten verbessern,
- der Einfluss dieser Massnahmen auf die User Experience evaluiert werden,
- der Auftraggeberin SBB Empfehlungen zum Relaunch des Ticketshops gegeben werden.

### 1.3 Abgrenzung

Die Arbeit beschränkt sich auf die Desktop-Version des Ticketshops. Die Mobile App ist abgeschlossen, da diese ein abweichendes Backend nutzt und somit nicht direkt mit der Desktop-Version vergleichbar ist. Auch haben wir aus praktischen Gründen darauf verzichtet, die Prototypen mit responsivem Design zu gestalten und haben nur eine Desktop-Version entwickelt, um so im Rahmen des gegebenen Aufwands zu bleiben.

Innerhalb des Ticketshops konzentrieren wir uns auf die Seite Reiseoptionen (Abbildung 1.2). Diese Seite ist zentral im neuen Ticketshop von SBB, wo der Benutzer alle wichtigen Daten zu seiner Reise eingibt. Wir haben uns für einen einfachen Anwendungsfall entschieden, damit die Umfrageteilnehmer die Aufgabe besser verstehen und nicht abgelenkt werden.

Zusätzliche Information zu Vorgehen, Design und Umsetzung der Tests sind ab Kapitel 3 zu finden.

### 1.4 Allgemeines Vorgehen

Unser Vorgehen ist durch den konkreten Fall des SBB-Ticketshops geprägt. Die Analyse des Systems und des Kontexts macht die Anwendung gängiger Methoden des Human Computer Interaction Designs (HCID) notwendig, wir setzen hier klassische Methoden wie Stakeholder Interviews und die Konkurrenzanalyse ein.

Das Hauptziel dieser Arbeit ist, aus diesem konkreten Fall allgemeine Erkenntnisse ableiten zu können, die die User-Experience-Gemeinde in ihrer Herangehensweise im Umgang mit dem La-  
deverhalten einer Applikation unterstützen kann.

Die Ziele der Masterarbeit sind allein mit qualitativen HCID-Methoden nicht zu erreichen, da diese nur schwer zu überprüfende Ergebnisse liefern und zu ungenau sind. Vielmehr ging unsere Planung in Richtung einer wissenschaftlichen Arbeit, um überprüfbare Erkenntnisse und Beweise zu liefern.

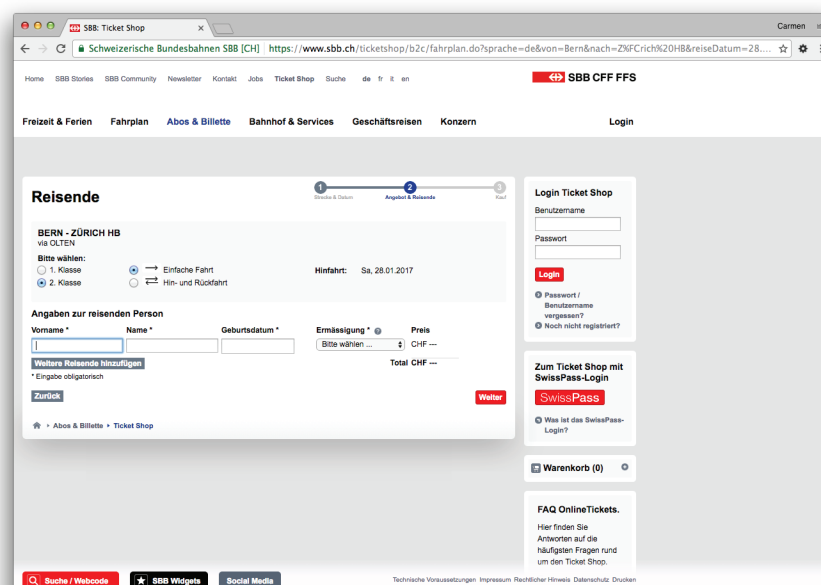


Abbildung 1.1: Aktuelle Webshop

Das Kernstück unserer Arbeit bildet daher die Durchführung eines Experiments und einer quantitativen Analyse. So durchlaufen wir die klassischen Schritte eines Experiments: Problemstellung, Hypothesenbildung, Versuchsplan, Versuchsaufbau, Versuchsdurchführung, Auswertung und zuletzt Interpretation. Trotz dieses Ansatzes bietet es sich an, in der Vorbereitung der Analyse und Durchführung zusätzlich qualitative Methoden ergänzend einzusetzen, um insgesamt bessere und sinnvollere Erkenntnisse zu gewinnen.

Ein grosser Teil der Masterarbeit besteht zunächst darin, einen groben Plan zu erstellen, die Aufgaben in verschiedene kleinere Abschnitte und Ziele aufzugliedern und den so entstandenen Plan Schritt für Schritt durchzuführen. Wir beginnen mit einer Literaturrecherche, die wichtige Erkenntnisse über das Thema Performance im Bereich UX und über die wissenschaftliche Vorgehensweise liefert. So kann die Problemstellung genauer definiert und erarbeitet werden.

Als nächstes definieren wir aufgrund dieser Recherche Hypothesen, die wir untersuchen. Gleichzeitig arbeiten wir bereits an den Merkma-

len und der Definition von Variablen, die durch unsere Studie bewiesen oder widerlegt werden sollen. Dieses Vorgehen ermöglichte uns, im Vorfeld des Experimentes die Hypothesen weiter zu ergänzen und zu erweitern, um sinnvollere Ergebnisse zu erzielen.

Indem wir unsere Hypothesen und Variablen iterativ erweitern, folgen wir einem experimentellen Versuchsplan. Im darauf folgenden Versuchsaufbau analysieren wir die zu operationalisierenden Variablen und entwickeln einen ersten Prototypen. Des Weiteren planen wir die Rekrutierung der Testpersonen. Die Durchführung des Experiments wird im Vorversuch durch qualitative Methoden, insbesondere Usability-Tests, ergänzt, um so die Richtigkeit des Vorgehens zu überprüfen und Störfaktoren möglichst auszuschliessen.

Als Letztes müssen die gesammelten Daten analysiert und interpretiert werden und Hypothesen und Theorien überprüft werden. Dazu wenden wir statistische Methoden an.

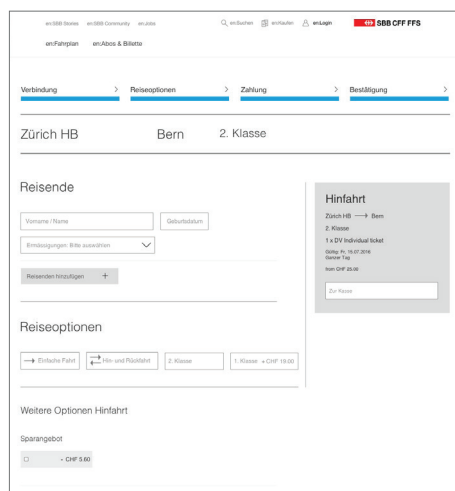


Abbildung 1.2: Wireframe der Reiseoptionenseite



## 2. Theoretischer Hintergrund

Anfangs stellt sich die Frage, wie Ladeverhalten und Performance überhaupt im Rahmen dieser Arbeit definiert werden können und welche Bedeutung sie für die User Experience haben. In diesem Kapitel beschreiben wir die theoretischen Grundlagen unserer Arbeit und wichtige Erkenntnisse aus der Literaturrecherche.

### 2.1 Performance

Die Ladezeiten von Websites und Webapplikationen zu verbessern, ist seit längerem eine Top-Priorität grosser Internet-Dienstleister wie Amazon, Facebook oder Google. Denn sie haben den Faktor *Ladezeit einer Webseite* als äusserst wichtig erkannt. Der Zusammenhang zwischen geringeren Ladezeiten und positiven Auswirkungen, wie längerer Nutzungsdauer, erhöhter Nutzungsfrequenz und grösserer Nutzerzufriedenheit, gilt heutzutage als erwiesen (Schurman/ Brutlag, 2009 and King, 2008). Bereits wenige Millisekunden können Auswirkungen auf die Geschäftsergebnisse haben (Souders, 2009).

Zudem trägt die gestiegene Erwartungshaltung der Benutzer dazu bei, dass Ladezeiten stets weiter optimiert werden müssen. Denn

- unabhängig vom Endgerät und Standort erwarten 47% der Konsumenten, dass eine Webseite innerhalb von zwei Sekunden geladen ist (Akamai, 2009).
- unabhängig von den mobilen Datenraten erwarten 71% der mobilen Benutzer, dass eine Webseite gleich schnell geladen wird wie auf dem Desktop (Bixby, 2011).

- 74% der Benutzer verlassen eine Webseite, wenn sie nicht innerhalb von 5 Sekunden geladen ist (Akamai, 2013).

Trotz der immer schneller gewordenen Internetverbindungen besteht das Problem auch heutzutage noch. Zwar haben die Down- und Upload-Geschwindigkeiten zugenommen, aber gleichzeitig sind auch die durchschnittlichen Datenmengen von Webseiten stark angestiegen. Beispielsweise ist die Datenmenge einer heutigen Webseite vergleichbar mit der eines kompletten 3D-Computerspiels vergangener Tage (Ronan, 2016). Da die Download-Geschwindigkeiten ähnlich rasch angestiegen sind wie die Datenmengen, haben sich die durchschnittlich benötigten Ladezeiten von Webseiten nicht sehr stark verändert (Belshe, 2010).

#### 2.1.1 Das Ladeverhalten im Kontext der ISO 9241-110 Richtlinien

Wie oben beschrieben, ist das Ladeverhalten von entscheidender Wichtigkeit für den Erfolg einer Applikation. Zusätzlich stellt sich die Frage nach der Bedeutung des gesamten Ladeverhaltens und der begleitenden Aspekte im Human Computer Interaction (HCI)-Forschungsfeld.

Einen ersten Orientierungspunkt stellt dabei der ISO 9241-110 Standard dar. Dieser Teil des ISO Standards handelt von allgemein etablierten Grundprinzipien, die bei dem Design von Dialogen zwischen Mensch und Informationssystemen zu beachten sind.

Ein Grundprinzip dieses Design-Standards ist die Rückmeldung des Systemzustandes an den Nutzer, so dass der Nutzer stets über den aktuellen Zustand des Systems informiert ist.

Der Fokus der Forschung im HCI- und Ergonomie-Bereich in Bezug auf Fortschrittsanzeigen und Ladeverhalten war in der Vergangenheit vor allem auf dieses Grundprinzip gerichtet (Lallemant/Gronier, 2012). Zum Beispiel bedeutet dies, dass die Forschung darauf abzielte, herauszufinden, inwieweit der Nutzer durch Ladebalcken und ähnliche Elemente imstande ist, Rückschlüsse auf den Ladezustand des Systems zu ziehen.

Dabei stellte sich interessanterweise heraus, dass gerade die Fokussierung der Aufmerksamkeit des Nutzers auf die Rückmeldung des Ladezustandes die wahrgenommene Ladezeit verlängert (Lallemant/Gronier, 2012). Dies bedeutet, dass nicht allein die Rückmeldung oder auch die reale Ladezeit entscheidend für eine positive Nutzererfahrung sein kann. Vielmehr stellt sich die Frage nach weiteren Einflüssen und deren Zusammenspiel.

### 2.1.2

## Wahrgenommene Performance

Es lohnt sich also, noch einmal genauer zu betrachten, was Ladezeiten und Performance überhaupt für die Benutzer bedeuten.

Abgesehen von technischen Möglichkeiten zur Verringerung der Ladezeit von verschiedenen Operationen und Bestandteilen einer Webapplikation, wie beispielsweise Bildern, HTML, Skripten usw., gilt das von den Benutzern selbst wahrgenommene Ladeverhalten als entscheidend.

Es beinhaltet unter anderem, ab wann Benutzer einen einzelnen Zustand einer Webseite als fertig geladen betrachten, ob sie eine Applikation insgesamt als performant erachten, oder ob sie nahtlos und flüssig mit der Applikation interagieren können.

Beispielsweise hat sich gezeigt, dass man die wahrgenommene Ladezeit verkürzen kann, indem man die Aufmerksamkeit der Benutzer auf andere Dinge lenkt. Daher besteht eine Möglichkeit, die wahrgenommene Ladezeit zu verkürzen, darin, Animationen zu verwenden oder aber Interaktionsmöglichkeiten bereitzustellen (Hohenstein, 2016). Wichtig ist dabei, den *Flow*, das heisst, den Gedankenstrom des Benutzers während der Interaktion mit einer Applikation, nicht zu unterbrechen und ihn möglichst nicht aus der Interaktion herauszureissen (Doherty, 2015). Andere Ansätze, die wahrgenommene Ladezeit zu optimieren, sind daher beispielsweise wichtige Teile einer Applikation zuerst zu laden und anzuzeigen, oder auch nur den Anschein zu erwecken, die Applikation sei bereits fertig geladen und bereit zur Interaktion, auch wenn der technische Ladevorgang noch nicht vollständig abgeschlossen ist.

## 2.2

### Responsiveness

Dieser Zusammenhang zwischen technischer Performance und der Wahrnehmung des Benutzers kann durch das Prinzip Responsiveness zusammengefasst werden, welches Jeff Johnson in seinem Buch GUI Bloopers 2.0 (2008: 45f.) beschreibt. Dieses erarbeitete er neben anderen Designprinzipien nach vielen Jahren Forschung im Bereich Lernen und Wahrnehmung von Menschen.

Er definiert Responsiveness als die Fähigkeit der Anwendungssoftware, mit den Benutzern Schritt zu halten («keep up with the user»). Die Benutzer sollen durch Feedback der Software wissen, was gerade geschieht, auch wenn die Software einen User Request nicht unmittelbar beantworten kann.

Responsiveness ist mit Performance verwandt, denn beide Begriffe haben mit Zeit zu tun. Performance wird mit der Systemzeit durch reine Zeiteinheiten gemessen, Responsiveness mit der Zeitwahrnehmung der Benutzer und letzten Endes mit ihrer Zufriedenheit. In diesem Sinn kann man Systeme finden, die sehr schnell sind, aber eine sehr schlechte Responsiveness haben, weil die Benutzer kein Feedback vom System bekommen. Auch das Gegenteil kann passieren, wenn ein langsames System fähig ist, die Benutzer immer über seinen Status und die Ausführung der Aktionen zu informieren.

Da Responsiveness von der subjektiven Zeitwahrnehmung der Benutzer abhängig ist, ist sie schwer zu messen und es lassen sich nur schwer konkrete Zeiten für adäquate Antwortzeiten vorgeben. Es existieren jedoch breit akzeptierte Richtlinien für Software-Antwortzeiten, die von

Card, Robertson und Mackinlay (1991: 181f.) entwickelt und von Nielsen (1993: 135) verbreitet wurden. Card und seine Mitautoren definierten bereits 1991 folgende Kategorien:

- 100 ms: «perceptual processing time constant». Wenn ein Prozess weniger als 100 ms dauert, nehmen die Benutzer ihn als unmittelbar ausgeführt wahr.
- 1 s: «immediate response time constant». Die Benutzer nehmen die Verzögerung zwar wahr, werten sie aber als unterbruchsfrei.
- 10 s: «unit task time constant», ist auch als Konzentrationsspanne bekannt. Dies ist die maximale Zeit, in der man die Aufmerksamkeit der Benutzer auf dem System halten kann.

Daraus folgend muss eine Operation den Benutzern kein Feedback geben, wenn sie weniger als 1 Sekunde dauert. Für alle Operationen, die länger dauern, empfehlen Card wie auch Nielsen, den Benutzern Feedback zu geben. Nielsen empfiehlt den Gebrauch von Fortschritts- und Ladeanzeigen, wenn eine Operation länger als 10 Sekunden dauert (1993: 186).

Steven C. Seow (2008: 49f.) bietet eine andere Möglichkeit zur benutzerzentrierten Bestimmung der Responsiveness.

Für Seow ist Responsiveness im Kontext der Interaktion zu betrachten und wird zusätzlich von den Benutzern subjektiv interpretiert. Er hat Zeitbereiche basierend auf der Erwartungshaltung der Benutzer definiert, um daraus eine akzeptable Antwortzeitspanne zu bestimmen. Er hat somit bewusst keine absoluten Zahlen bestimmt, wie das Cards Modell tut. Die Antwortzeiten werden von einer von den Benutzern aktiv gestarteten Aktion bis zu der von ihnen wahrnehmbaren Systemantwort berechnet. Seow unterscheidet in seinem Modell vier Typen von Zeitspannen (Abbildung 2.1):

**Instantaneous** (0.1 to 0.2 s): zwischen 100 und 200 Millisekunden. Die Benutzer erwarten von Objekten oder Gegenständen in der virtuellen Welt (z.B. Button) grundsätzlich dasselbe Verhalten und dieselbe Geschwindigkeit, wie sie es von der realen Welt (z.B. einem analogen Lichtschalter) kennen: im Prinzip müssen Dinge sofort passieren (instantaneous).

**Immediate** (0.5 to 1 s): So, wie Benutzer den Austausch von Informationen zwischen Menschen erleben (Frage - Pause - Antwort), erwarten sie auch Feedback oder *responsiveness* auf ihre Aktionen in der virtuellen Welt. Sie wollen unmittelbar wissen, dass die ausgelöste Aktion verarbeitet wird. Seow spricht hier von einer halben bis zu einer ganzen Sekunde Wartezeit. Ein

Beispiel ist das Klicken auf einen internen Link einer Webseite, bei dem Benutzer eine leichte Verzögerung akzeptieren.

**Continuous** (2 to 5 s): Die Benutzer verstehen, dass das System für eine Operation länger braucht und sie bekommen vom System entsprechendes Feedback, damit die Interaktion zwischen ihnen und dem System nicht abreißt. Das System sollte die Benutzer folglich bereits informieren, wenn eine Antwortzeit zwischen 2 und 5 Sekunden benötigt wird.

**Captive** (7 to 10 s): Dieser Zeitspannentyp stimmt mit Cards *unit task time* überein. Bei höheren Wartezeiten wird ein System Schwierigkeiten haben, die Aufmerksamkeit der Benutzer zu halten. Also muss das System den Benutzern Feedback geben und sollte ihnen gleichzeitig die Freiheit geben, die Zeit für andere Zwecke zu nutzen (Miller, 1968: 267f.)

Fiona Nah (2004: 153f.) zeigt, dass Benutzer im Internet nicht gerne warten und nach 2 Sekunden Wartezeit beginnen, das Interesse zu verlieren. Benutzer sind zufriedener, wenn sie vom System Feedback bekommen. Weiter sagt die Studie von Branaghan und Sanchez (2009: 528f.) aus, dass Feedback-Elemente, die die Aufmerksamkeit der Benutzer zu sehr auf sich ziehen, gleichzeitig verursachen, dass die Wartezeit als länger

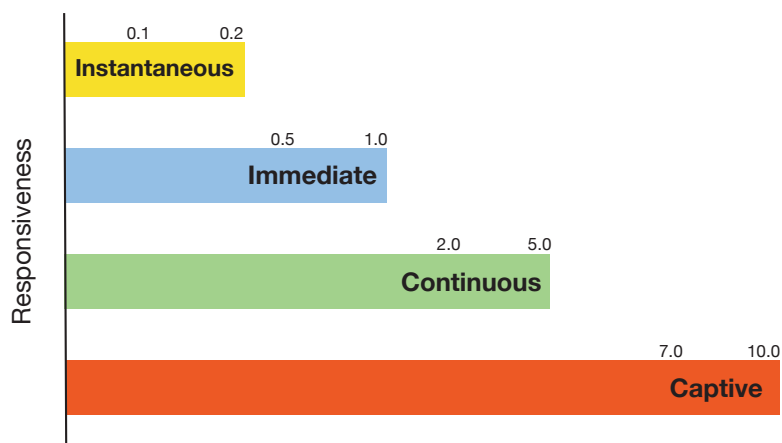


Abbildung 2.1: Typen von Zeitspannen(Seow, 2008)

eingeschätzt und somit als weniger angemessen betrachtet wird. Trotzdem bevorzugen die Benutzer es, Feedback zu bekommen, auch wenn die Wartezeit durch das Feedback länger erscheint. Es ist also ein Balanceakt, die Faktoren Zeit, Feedback und Erwartungshaltung optimal aufeinander abzustimmen.

### 2.2.1 Fortschrittsanzeigen

*Progress indicators* oder Fortschrittsanzeigen (auch Ladevorgangsanzeigen) sind eine der meistgenutzten Lösungen für Feedback während der Ladezeit. Sie geben den Benutzern einerseits Feedback, belasten sie andererseits aber auch nicht mit zu viel zusätzlicher Information. Zudem erfüllen Fortschrittsanzeigen Maisters Aussagen aus *The Psychology of Waiting Lines* (1985: 113f.), im Speziellen die Aussagen «Ungewisse Wartezeiten sind länger als bekannte, endliche Wartezeiten» und «Unerklärte Wartezeiten sind länger als erklärte Wartezeiten» (Mishunovs, 2015c).

Seow zeigt in seinem Buch einen Weg auf, wie Fortschrittsanzeigen klassifiziert werden können und bietet eine Hilfestellung, wie man für eine bestimmte Anwendung und Wartezeit die richtige Anzeige bestimmen kann. Seow (2008: 80f.)

unterscheidet zwei Eigenschaften der Indikatoren: bestimmt/unbestimmt (weiss man, wann der Prozess zu Ende ist?) und dynamisch/statisch (Gibt es Informationen während der Prozess läuft?). Die Kombination dieser Eigenschaften ergibt vier Typen von Fortschrittsanzeigen (Abbildung 2.2).

**Class A(ctive):** Bei dieser Anzeige wissen die Benutzer, wann der Prozess zu Ende sein wird und werden laufend über den Status informiert (bestimmt/dynamisch). Das bekannteste Beispiel ist der Fortschrittsbalken.

**Class D(isappear):** In diesem Fall werden die Benutzer direkt über das Ende des Prozesses informiert. Dieser läuft dann aber im Hintergrund ohne weiteres Feedback und lässt den Benutzern so die Freiheit, währenddessen eine andere Tätigkeit zu verrichten (bestimmt/statisch). Ein gutes Beispiel ist ein sehr langer Downloadprozess.

**Class C(hanges):** Wenn es nicht möglich ist, den Benutzern den Endzeitpunkt eines Prozesses anzuzeigen, kann der Fortschritt auch in anderen Einheiten angezeigt werden. Ein Antiviren-Programm zeigt zum Beispiel, wie viele Dateien bereits gescannt wurden (unbestimmt/dynamisch).

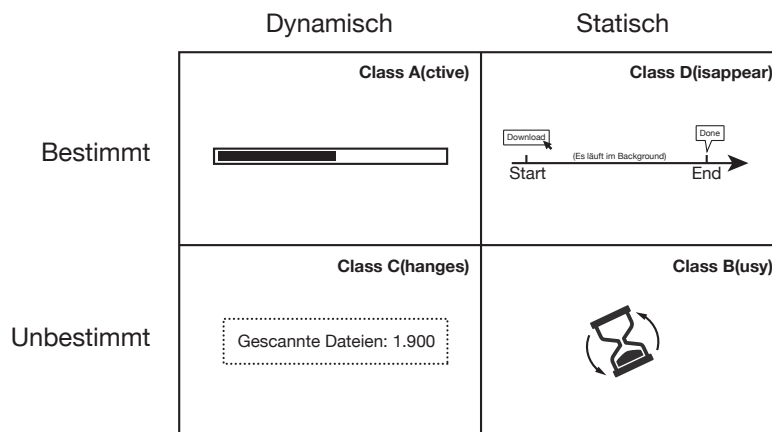


Abbildung 2.2: Übersicht von Fortschrittsanzeigen (Seow, 2008)

**Class B(usy):** Dieser Indikator gibt den Benutzern keine Information über die Dauer, sondern zeigt nur an, dass das System arbeitet (unbestimmt/statisch). Sehr bekannt und verbreitet in dieser Klasse ist die drehende Sanduhr von Windows oder das drehende Rad (Spinning Wheel, auch Spinner genannt) in einer Webapplikation (Abbildung 2.3).

Die Empfehlung für die Benutzung einer Fortschrittsanzeige bezieht sich auf die Zeitspanne der Responsiveness (Seow, 2008: 80f.):

- Für *instantaneous*-Reaktionszeiten wird **keine Anzeige** benötigt.
- Auch bis zu einer Dauer von zwei Sekunden (*Immediate*) ist eine Anzeige nicht nötig, es schadet aber nicht, **einen einfachen Indikator ohne Text (Class B)** zu zeigen.
- Ab zwei Sekunden (*Continuous*) ist es nötig, den Benutzern Feedback in Form einer Fortschrittsanzeige zu geben. Bis zu fünf Sekunden genügt normalerweise die Verwendung eines *Class B*-Indikators mit **einer typischen Busy-Anzeige** (z. B. einem drehenden Rad).
- Wenn die Wartezeit über 5 Sekunden liegt (*Continuous*), wird **ein Indikator des Typs A oder C** empfohlen.

- Für Prozesse, die mehr als 10 Sekunden brauchen, sollte man den Benutzern zusätzlich die Möglichkeit geben, den Prozess abbrechen oder verlassen zu können.

Im online-Magazin *Smashing Magazine* kann man neuere Richtlinien finden, die sich eher auf Webseiten beziehen. Diese Richtlinien stimmen mit Seows Empfehlungen teilweise überein. Denys Mishunovs (2015c) definiert als Vorgaben für Webseiten, dass bei einer Ladezeit von:

- 0.5 bis 1s die Notwendigkeit einer Anzeige fallweise entschieden werden sollte.
- 1 bis 5s ein *Class B*-Indikator (drehendes Rad) oder ein einfacher *Class A*-Indikator (Ladebalken ohne Text) verwendet werden sollte.
- mehr als 5s ein dynamischer Indikator, entweder ein *Class A*-Indikator mit dynamischen Texten oder Prozenten, oder ein *Class C*-Indikator mit für die Benutzer wichtigen Informationen über den Prozess verwendet werden sollte.

Was heutzutage auch im Internet zu finden ist, sind Kombinationen von verschiedenen Anzeigeklassen. Beispielsweise benutzen Slack und Outlook dynamische *Class B*-Indikatoren, die



Abbildung 2.3: Beispiele von Spinners

nach kurzer Zeit zusätzliche Informationen über den Status des Prozesses geben (Abbildung 2.4). Der Endzeitpunkt des Prozesses ist nicht bekannt, aber die Benutzer werden regelmässig informiert.

Bei unserer Analyse von Konkurrenzseiten (siehe Kapitel 3.2.2) aus anderen Ländern und von Seiten mit ähnlichen Services (z.B. Fluggesellschaften) beobachteten wir, dass die meistgenutzte Anzeige ein einfacher *Class B*-Indikator mit einem drehenden Rad (Spinner) ist. Obwohl wir die Wartezeiten nicht genau messen konnten, befand sich der grösste Teil über der 5 Sekunden-Marke. Somit halten sich diese Seiten nicht an die oben beschriebenen Richtlinien oder an aktuelle Aussagen von bekannten Persönlichkeiten aus dem UX-Umfeld, wie Adrian Zumbrunnen oder Luke Wroblewski. Beide empfehlen ausdrücklich, den Spinner nur zurückhaltend zu benutzen (Zumbrunnen, 2015) oder sogar komplett darauf zu verzichten (Wroblewski, 2013 und 2014).

*«While the intentions behind progress indicators are good, the end result can actually turn out to be bad because by definition they call attention to the fact that someone needs to wait».*

– Luke Wroblewski

### 2.3

## Visuelle Ästhetik des Ladeverhaltens

Die Wahrnehmung von Ladeverhalten und Fortschrittsanzeigen geht über eine rein nüchterne Rückmeldungsfunktionen hinaus. Wie andere Teile einer Webapplikation auch, wird das Ladeverhalten durch die Benutzer nach ästhetischen Gesichtspunkten beurteilt. Es beeinflusst somit die gesamte Zufriedenheit der Benutzer und die User Experience der Applikation.

Nach Norman geschieht die Wahrnehmung und Verarbeitung von Informationen beim Menschen auf drei Ebenen: visceral, behavioral und reflective. Jede dieser Ebenen bezeichnet eine andere Art, wie Menschen Informationen auf eine positive oder negative Weise bewusst und unterbewusst wahrnehmen (Norman, 2004: 3). So können ästhetische Gestaltung und attraktive Darstellungen auf diesen Ebenen unmittelbare Reflexe hervorrufen (-> *visceral*), antrainierte Assoziationen wecken (-> *behavioral*) oder auch zum Nachdenken anregen (-> *reflective*).

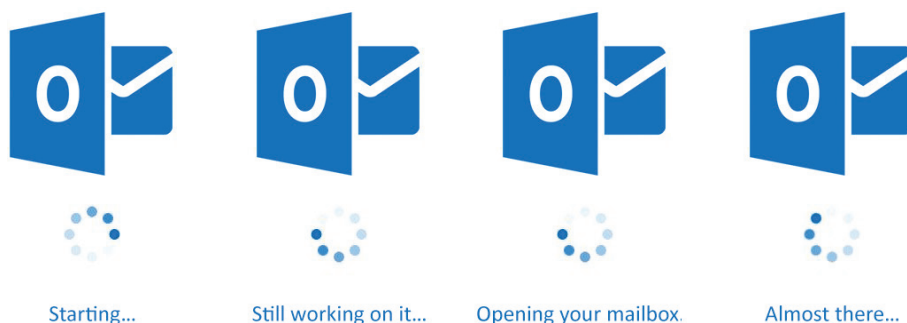


Abbildung 2.4: Fortschrittsanzeige Outlook

Besonders unmittelbare und antrainierte Reaktionen und Assoziationen laufen innerhalb weniger Sekunden ab – also innerhalb der Zeit, die ein Ladevorgang dauern kann – und sind daher für diese Arbeit entscheidend. Aus diesem Grund können Fortschrittsanzeigen auch bei Applikationen, deren sonstiges Auftreten eher seriös und konservativ gehalten ist, die Möglichkeit bieten, durch kleine Animationen oder Illustrationen für einen emotionalen Überraschungseffekt zu sorgen.

Das bedeutet, dass Fortschrittsanzeigen eine Gelegenheit bieten, die rein funktionale Optimierung der Usability weiterzuführen und mit Emotionen eine stärkere Bindung zu der Applikation zu schaffen. Denn eines ist erwiesen: Attraktive Dinge funktionieren besser (Norman, 2004: 17).

## 2.4

### **Bedeutung des theoretischen Hintergrunds für unsere Arbeit**

Wie oben beschrieben ist ein erfolgreiches Ladeverhalten von vielen Faktoren abhängig. Für unsere Arbeit folgt daraus, dass wir die entscheidenden Faktoren und ihr Zusammenspiel im Kontext des konkreten Anwendungsfalls definieren und untersuchen müssen.

Auch wenn die theoretischen Grundlagen bekannt sind, müssen wir erfolgreiches Ladeverhalten im Fokus dieser Arbeit genau definieren und operationalisieren. Dies soll im besonderen Masse unter Bezug auf die User Experience geschehen. Das heisst, dass sowohl die wahrgenommene Ladedauer möglichst gering sein soll, als auch die Zufriedenheit der Benutzer mit dem Ladeverhalten möglichst hoch sein soll.

Das entscheidende Mittel in diesem Bereich ist dabei die Rückmeldung der Applikation über ihren Zustand der Applikation an den Benutzer. Dieses kann durch vielerlei Dinge, wie progressives Ladeverhalten oder Priorisierung der Ladeihenfolge von Teilen der Applikation optimiert werden. Im Zentrum dieser Arbeit steht jedoch der Ladeindikator als wichtigstes und meistgenutztes Mittel zur Informationsübermittlung und Rückmeldung.

Des Weiteren konzentrieren wir uns im Folgenden auf ausgewählte wichtige Merkmale und Faktoren des Ladeverhaltens, um unsere Untersuchungen weiter einschränken und fokussieren zu können. Dadurch erhoffen wir uns aussagekräftige und sinnvolle Ergebnisse.

Der erste entscheidende Faktor ist die Ladezeit als Kernproblem des SBB-Ticketshops. Die Untersuchungen und Analysen der Arbeit befassen sich daher mit der Zeitspanne des Ladevorgangs. Hier gibt es aufgrund der vorhandenen Theorie zwei Dinge zu beachten. Einerseits muss nicht die eigentliche Ladezeit, sondern die vom Benutzer wahrgenommene Ladezeit genauer definiert und untersucht werden. Zum anderen wird die wahrgenommene Zeit durch unterschiedliches Ladeverhalten beeinflusst, was es bei der Operationalisierung der zu untersuchenden Variablen zu beachten gilt.

Der zweite zu untersuchende Faktor sind die Fortschrittsanzeigen selbst. Sie sind das wichtigste Mittel zur Rückmeldung des Ladezustandes. Dabei gibt es durch die untersuchte Theorie Hinweise, dass je nach Anwendungsfall und Ladezeit unterschiedliche Typen von Fortschrittsanzeigen optimal sind. Darüber hinaus wird zu untersuchen sein, inwiefern diese unterschiedlichen Typen von Fortschrittsanzeigen die Zufriedenheit der Benutzer mit dem Ladeverhalten beeinflussen. Die Position des Ladeindikators auf dem Bildschirm soll als weiterer Faktor im Auge behalten werden.

Als drittes ist die Ästhetik und Gestaltung des Ladeverhaltens wichtig für die Gesamtbeurteilung der User Experience und die Benutzerzufriedenheit. Besonders im Bereich von wenigen Sekunden bietet eine überraschende oder ungewohnt andersartig gestaltete Ladeanzeige eine Möglichkeit, positive Emotionen bei den Benutzern hervorzurufen.



# 3. Vorbereitung der Studie

Nachdem die Bedeutung der Theorie in Bezug auf das Ladeverhalten klar geworden ist, ist es notwendig, die ersten Analyseschritte am konkreten Fall in den HCID-Kontext einzuordnen. Dabei dienen als Vorbereitung zur eigentlichen wissenschaftlichen Studie Methoden und Prinzipien des allgemeinen HCID-Vorgehens. Beim eigentlichen Experiment werden wir auf das allgemeine wissenschaftliche Vorgehen zurückgreifen.

## 3.1

### User-Centered-Design Prinzipien

Wir orientieren uns zuallererst an allgemeinen User-Centered-Design-Prinzipien, um das Verständnis für die Problemstellung aus Nutzersicht zu fördern. Dies hat den Vorteil, dass unsere Hypothesen und Experimente mehr Aussagekraft bekommen. Dies sind unter anderem die folgenden Prinzipien (Gulliksen et al., 2003: 397f.):

- Das Design basiert auf einem expliziten Verständnis von Benutzern, Aufgaben und Kontext.

Es ist wichtig und Teil dieser Arbeit, vor allem das System und den Kontext zu verstehen, daher werden einige der einschlägigen HCID-Methoden zur Analyse verwendet.

- Benutzer sind in den Design- und Entwicklungsprozess involviert.

Beim Vorversuch unserer wissenschaftlichen Studie planen wir, Benutzer bereits einzubeziehen, um so einerseits unsere Hypothesen besser entwickeln zu können und andererseits einen reibungslosen Ablauf des Experiments sicherzustellen.

- Das Design wird durch benutzerzentrierte Evaluation vorangetrieben.

Methoden wie Usability-Tests, Hallway Tests und Fragebogen stellen einen direkten Bezug zu den Benutzern dar und gewährleisten fortwährende valide Erkenntnisse.

- Der Prozess wird iterativ durchgeführt. Vor allem unsere Hypothesen werden im Prozess der Studienvorbereitung fortlaufend erneuert, erweitert und ausgebaut.

- Im Design-Team sind multidisziplinäre Fähigkeiten und Perspektiven vorhanden.

Die Bandbreite an Know-How reicht im Team von Programmieren bis zur HCID Methodik und visuellem Design. Das ermöglicht es, neue Perspektiven zu bekommen und Hypothesen besser zu entwickeln.

Um die Problemstellung und den Kontext genauer zu verstehen, bietet sich als erster Schritt an, den Kontext des Systems zu analysieren und spezifizieren.

## 3.2.

### Analyse und Spezifikation des Kontexts des Systems

In diesem Schritt wenden wir einige konkrete HCID-Methoden an. Diese Methoden dienen der Analyse und Spezifikation des Kontexts des Systems, ähnlich der ersten Aktion des einschlägig bekannten ISO-Modell 9241-210, welches ein Standardvorgehen des User Centered Designs beschreibt. Insbesondere das Stakeholder Interview und die Konkurrenzanalyse erscheinen uns

besonders geeignet, um einen Überblick über die Sachlage zu bekommen. Wie die Literaturrecherche (siehe Kapitel 2) sind diese Methoden hilfreich, Informationen zu sammeln und das Verständnis zu verbessern. Dies dient insbesondere als Vorbereitung für die darauf folgende Phase, in der wir die wissenschaftliche Methode anwenden.

### 3.2.1 Stakeholder Interview

Das Stakeholder Interview ist eine der ersten Methoden, die angewendet werden sollen, damit das Projektteam wichtige Informationen des gesamten Kontexts sammeln kann und damit das Projekt ohne spätere Überraschungen ablaufen kann. Das Interview dient ausserdem auch dazu, den/die Stakeholder besser kennenzulernen. Die genaue Motivation hinter dem Projekt aufzufindig zu machen, hilft beim weiteren Vorgehen. Zum Beispiel können Aspekte priorisiert oder eliminiert werden, wenn die Auftraggeber sie für wichtig oder unwichtig erachten. Ein Stakeholder Interview gibt dem Projektteam die Plattform, gezielt Fragen zu stellen, zum Beispiel nach der Wichtigkeit des Projekts für den Auftraggeber, den Zeitrahmen für das Ausrollen der Ap-

plikation zu verstehen und herauszufinden, wie der technische Rahmen aussieht und was die technischen Abhängigkeiten für dieses Projekt sind. Am allerwichtigsten ist es, zu wissen, wer die Benutzer sind und für wen diese Applikation gebaut wird. Je nach Organisation des Unternehmens, das den Auftrag vergibt, variieren die Stakeholder. Ein Projektleiter, ein technischer Leiter oder ein Executive Director sind Kandidaten, die befragt werden können. Abgesehen von seinem Titel muss ein Stakeholder in erster Linie auf das Projekt selber Einfluss nehmen können. Beispiele dafür, wie ein Auftraggeber auf das Projekt Einfluss nehmen kann, sind die Finanzierung und die Möglichkeiten, die Applikation testen zu lassen und die Applikation zu konstruieren (Goodwin, 2009: 28f.)

Unser Stakeholder für dieses Projekt ist Mike Morgenthaler, der UX-Lead der SBB (Abbildung 3.1).

Um das Thema und die SBB-Umgebung besser zu verstehen, reisten wir nach Bern um uns mit Mike Morgenthaler zu treffen, dessen Idee es gewesen war, den Ladevorgang genauer zu untersuchen. In diesem ersten Interview stellten



Abbildung 3.1: Foto von Mike Morgenthaler

wir die wichtige Frage, weswegen die SBB das Thema Fortschrittsanzeige untersuchen wollten. Morgenthaler berichtete uns, dass auf Basis der Erfahrungen der Projektleiter die Idee entstand, ein neues Konzept zu entwickeln, um den Prozess für die Benutzer zu optimieren. Jedoch wurde bisher keine User Research betrieben, da das UX-Team erst zu einem späteren Zeitpunkt gegründet wurde.

Der Prozess im neuen Webshop stellt vermehrt die Bedürfnisse der Benutzer in den Vordergrund. Das heisst konkret, dass der Benutzer angibt, wie er reisen möchte und dass der Webshop automatisch das Billet auswählt, das die Bedingungen am günstigsten erfüllt. Im alten Webshop muss der Benutzer dagegen manuell zwischen alternativen Billet-Angeboten wählen und sich so aus einer Vielzahl von Möglichkeiten das Endresultat selber zusammenklicken. So kann er nicht automatisch von Spezial-Angeboten profitieren. Der neue Webshop stellt durch die automatischen Berechnungen dagegen wiederum grössere Anforderungen an die Performance und an die User Experience.

Morgenthaler berichtete über den Aufbau der SBB-Applikation und die verschiedenen Backend-Abhängigkeiten (Abbildung 3.2). Das Angebot wird aus Daten unterschiedlicher Anbieter zusammengestellt. Durch diese Abhängigkeiten entstehen Zeitverzögerungen in den Ladezeiten, für die es in nächster Zeit keine technische Lösung geben wird. Für die SBB bedeutet dies, dass sie nach dem Rollout der Applikation mit Performance-Problemen rechnen müssen. Für unsere Arbeit bedeutet dies wiederum, dass wir keinen Einfluss auf die Zeitverzögerung nehmen können, die das neue Konzept mit sich bringt und unser Konzept danach aufbauen müssen.

Weitere Anforderungen und Einschränkungen sind, dass die Applikation sich nur auf eine Desktop-Version beschränkt, da die Mobile App noch das alte Backend-Konzept benutzt.

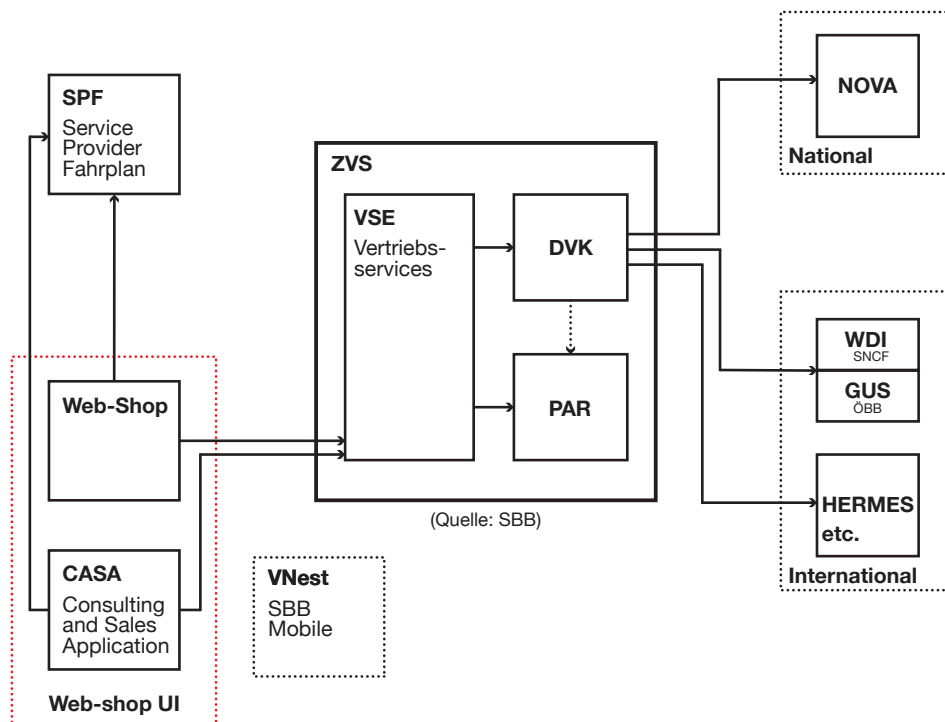


Abbildung 3.2: Backend der SBB zeigen

Das alte und das neue Backend verwenden inkompatible Anfragen / Antworten (API, Application Programming Interfaces). Das heisst, das neue Backend wurde für das neue Frontend designed. Das neue Frontend sendet andere Arten von Anfragen an das neue Backend. Die Konsequenz ist, dass das alte Frontend (die mobile App) nicht mit dem neuen Backend kommunizieren kann.

Auf unsere Rückfrage per E-Mail erklärte Morgenthaler uns, dass die Ladezeiten je nach Auslastung teilweise sehr unterschiedlich sind.

Die Durchschnittswerte liegen zwischen 4.8 und 8 Sekunden. Im Einverständnis mit unserem Stakeholder legen wir den Durchschnittswert auf 7 Sekunden (siehe Kapitel 5.2) fest. Weitere Anforderungen sind, dass wir keinen Einfluss auf das Backend ausüben können und dass die mobile Lösung ein anderes Backend benutzt.

Dadurch können wir eine klare Abgrenzung zur mobilen Seite schaffen. Diese Erkenntnisse sind erste Parameter für unser späteres Design.

Im Rahmen der Benutzer-Analyse der SBB wurden keine Benutzer-Gruppen befragt. Es muss hier erwähnt werden, dass es zu diesem Zeitpunkt das UX-Team innerhalb der SBB noch nicht gab. Die SBB haben dennoch eine Anzahl von Personas aufgestellt und daraus ein Panel von Benutzern zusammengestellt, die in der Pilotphase die neue Desktop-Applikation benutzen werden. Es handelt sich dabei um 500 Nutzer aus dem Bereich Zürich, die während 2-3 Monaten die Desktop-Applikation testen werden.

Wir erhalten die Merkmale dieser Personas von den SBB und benutzen einige davon bei der Rekrutierung der Testpersonen für die Usability-Tests. Diese sind:

- Wohnort: Bern oder Zürich
- Abonnement-Typ: kein GA-Abonnement vorhanden
- Kauf-Plattform: Benutzer von SBBs Webshops

Wir fragten an, ob wir das Benutzer-Panel der SBB nutzen dürfen. Dies würde es uns erleichtern, Benutzer zu finden, die zu diesem Thema passen. Wir können das Panel der SBB jedoch nicht benutzen. Wir fokussierten uns aber, gemäss der oben genannten Merkmale, für unseren Usability-Test auf Benutzer ohne GA, da deren Motivation grösser ist, den Desktop Shop (sbb.ch) zu benutzen (siehe Kapitel 6.2.2). Der grösste Absatzkanal für die SBB, so Morgenthaler, sind die Automaten, der zweitgrösste die mobile App und zum Schluss die Desktop-Variante.

Morgenthaler zeigte uns, wie die Applikation aussieht. Da sich zum Zeitpunkt unserer Masterarbeit die externe Design-Firma im Iterationsprozess befindet, dürfen wir das Design nicht eins zu eins übernehmen. Dies bedeutet für uns, dass wir das GUI des Prototypen ähnlich dem Design der SBB nachbauen. Der Wiedererkennungswert des GUIs ist wichtig. Der Benutzer soll unsere Prototypen als Seite der SBB wiedererkennen.

Schon früh entscheiden wir, dass nicht der ganze Einkaufsprozess dargestellt werden soll, sondern der Benutzer sich auf die Shop-Seite und deren Fortschrittsanzeige konzentrieren können soll. Während unseres Usability-Tests kann der Benutzer die Informationen aus vorgängigen Etappen aus dem beschriebenen Szenario entnehmen (siehe Kapitel 6.2.2).

Dieses erste Stakeholder Interview mit Mike Morgenthaler inspiriert uns, die im Kapitel 2.4 aufgestellte Hypothese weiterzuentwickeln.

### **Positionierung der Fortschrittsanzeige**

Wir stellen uns die Frage, ob die ganze Seite für den Ladevorgang blockiert werden muss oder ob man die Positionierung auf den eigentlichen Bereich reduzieren kann. Dieses Thema möchten wir vertiefen und herausfinden, was der "State of the Art" ist.

### **Bildliche Darstellung der Fortschrittsanzeige**

Eine weitere Idee, die diesem Interview entspringt, ist die bildliche Darstellung eines Zugs als Fortschrittsanzeige. Haben andere Anbieter schon daran gedacht, dies so zu gestalten? Um den erwähnten Fragen nachzugehen und das Thema zu vertiefen, führen wir eine Konkurrenzanalyse durch (siehe Kapitel 3.2.2).

### **3.2.2**

## **Konkurrenzanalyse**

Eine Konkurrenzanalyse artverwandter Seiten mit ähnlichem Ladeverhalten bietet sich an dieser Stelle an.

Eine Konkurrenzanalyse dient dazu, zu verstehen, welche Lösungen am Markt zu diesem Thema bereits erarbeitet wurden und dazu, die positiven und negativen Aspekte dieser Lösungen herauszuarbeiten. Ebenso hilft diese Art der Analyse, eine neue Lösung zu entwickeln, die sich von der Konkurrenz abhebt (Goodwin, 2009: 197f.).

Während unseres Stakeholder Interviews hat der UX-Lead Mike Morgenthaler unsere Vermutung bestätigt, dass das neue Konzept des SBB-Ticketshops von den Fluggesellschaften inspiriert ist. Um die aktuellen Standards der Industrie bzw. des Sektors besser kennenzulernen und einen Startpunkt für unsere Arbeit zu finden, haben wir die Webseiten der Konkurrenz und andere für das Thema relevante Seiten angeschaut.

Wir suchten auf thematisch ähnlichen Webseiten Beispiele von Fortschrittssanzeigen und untersuchten deren jeweilige besondere Verwendung. Wir analysieren neben Fluggesellschaften auch das Shop-Konzept von ausländischen Bahnen und deren Lösung für die Fortschrittsanzeige.

Insbesondere gilt es für den Bereich von ca. 5-10 Sekunden Ladedauer, Lösungen zu finden, die ähnlich unserer Haupthypothese (dass eine Fortschrittsanzeige die wahrgenommene Wartezeit verkürzt) umgesetzt waren. Dabei klassifizieren wir die gefundenen Fortschrittsanzeigen nach Seow wie im Kapitel 2.2.1 beschrieben (Abbildung 3.3), um eine bessere Vergleichbarkeit herzustellen.

Dabei stellt sich heraus, dass viele Webseiten mit Ladezeiten von 5-10 Sekunden den Ladeindikator nur über einen Teil des Bildschirms legen. Bei anderen Varianten deckt der Ladeindikator entweder den ganzen Bildschirm ab oder wird auf einem sehr beschränkten Bereich positioniert, wie z.B. auf einem Button oder einem Eingabefeld, das gerade geladen wird.

Die Beispiele zeigen uns zwei Dinge:

#### **Position des Ladeindikators**

Zum einen stellen Ladeanzeigen in einem Teilbereich einer Seite das Standardvorgehen bei Flug- und Reiseportalen mit längerer Wartezeit dar. Die Gründe mögen teils technischer Natur sein, aber sicherlich lässt sich auch absichtliches Design mit dem Ziel besserer User Experience als Grund vermuten. Wie bereits beschrieben, könnte das wahrgenommene Ladeverhalten wiederum entscheidend sein. Progressives Ladeverhalten mit einem schrittweisen Aufbau von Elementen ist wichtig, um die gefühlte Wartezeit zu überbrücken. Somit sind Fortschrittsanzeigen, die nur einem Teilbereich des User Interfaces zugeordnet sind, Ladeanzeigen vorzuziehen, die das gesamte Benutzer-Interface blockieren.

#### **Kreativität in der Fortschrittsanzeige**

Zum anderen lässt sich ein Streben nach einer gewissen Kreativität und Differenzierung bzw. Abwechslungsreichtum bei den Fortschrittsanzeigen feststellen. Die eingesetzte Ästhetik und Kreativität im Design von Fortschrittsanzeigen kann verschiedene Bedeutungen haben. Eine positive Auswirkung auf die User Experience lässt sich aber durchaus vermuten sowie theoretisch begründen (siehe Kapitel 2).

Für ein wissenschaftliches Experiment ergeben sich daher folgende Schlüsse:

**Position der Ladeanzeige:** Fast alle Seiten und Portale benutzen eine oder mehrere lokal auf einen bestimmten Bereich begrenzte Fortschrittsanzeigen statt einer zentralen Fortschrittsanzeigen für den ganzen Bildschirm.

Die Position der Ladeanzeige haben wir bisher in unserer theoretischen Betrachtung nicht beachtet, werden nun aber durch die Konkurrenzanalyse darauf gebracht, dies noch einmal im Detail zu betrachten und zu untersuchen. Dabei bietet es sich an, auch in unserem Experiment die Position des Ladeindikators auf einen Bereich zu beschränken. Störfaktoren, wie eine wechselnde Position oder unterschiedliches Aussehen der Fortschrittsanzeige je nach Variante können so vermieden werden.

**Kreativität bei der Ladeanzeige:** Es gibt bei einigen Portalen die Tendenz, die Fortschrittsanzeige etwas kreativer und ungewöhnlicher zu gestalten. Zudem stützt die vorhandene Literatur die These, dass durch visuelle Ästhetik die Zufriedenheit in der Benutzung durch eine Vielzahl von Emotionen, die auf verschiedenen Ebenen verarbeitet werden, positiv beeinflusst werden kann. Daher bietet es sich an, in einem Experiment eine Variante mit einer etwas ungewöhnlichen, kreativen Variante anzubieten und zu überprüfen.

**Skyscanner**  
skyscanner.net

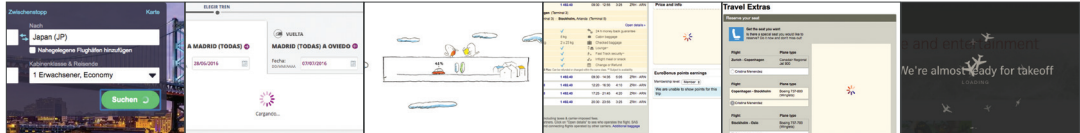
**Renfe**  
renfe.com

**Little Big Room SAS**  
little-big-room.com flysas.com

**SAS**  
flysas.com

**Swiss**  
swiss.ch

**Bild**



<b>Fortschritts-anzeige</b>	Busy-Anzeige: drehendes Rad (Class B, Seow, 2008)	Busy Anzeige mit Text (Class C: Mishunov, 2015c)	Fortschrittsbalken in spielerischer Art (Class A, Seow, 2008)	Busy-Anzeige: drehendes Rad (Class B, Seow, 2008)	Busy-Anzeige: drehendes Rad (Class B, Seow, 2008)	Busy Anzeige mit Text (Class C: Mishunov, 2015c)
<b>Position</b>	Auf dem Suchbutton	Auf dem Bereich, in dem die Änderungen stattfinden	Auf der ganzen Seite	Auf dem Warenkorb	Auf dem Bereich, in dem die Änderungen stattfinden	Auf der ganzen Seite

Abbildung 3.3: Fortschrittsanzeige der Konkurrenz (Brown, 2007: 113.)



# 4. Stadien der wissenschaftlichen Methode

Nach der Bestimmung des für unsere Arbeit relevanten Kontexts des Systems und der groben Ideenskizzierung für Hypothesen, planen wir im Rahmen der wissenschaftlichen Methode zu diesem Zeitpunkt die verschiedenen Stadien, die das Experiment durchlaufen muss.

## Wissenschaftliche Methode

Die Untersuchung der verschiedenen Aspekte des Ladevorgangs ist sehr vielschichtig (siehe Kapitel 1.4). Der Ladevorgang geschieht innerhalb von wenigen Sekunden, die Unterschiede der Varianten sind auf den ersten Blick minimal und eine Vielzahl von Variablen aus dem Kontext beeinflussen wiederum das wahrgenommene Ladeverhalten. Darüber hinaus geschieht Wahrnehmung zu einem bestimmten Teil unbewusst, das heißt, die beobachteten äusserlichen Reaktionen sind nur Anzeichen für tiefergehende psychologische Abläufe. Daher lassen sich aus unmittelbaren Beobachtungen nur schwer korrekte Schlussfolgerungen ziehen.

Erkenntnisse über die zu untersuchenden Variablen sind nur schwer aus rein qualitativen Untersuchungen (Usability-Test) zu gewinnen. Um wirklich sinnvolle Erkenntnisse zu gewinnen, ist die Durchführung eines wissenschaftliches Experiments notwendig und sinnvoll.

Diese wissenschaftliche Methode durchläuft diese sechs Stadien (Abbildung 4.1):

- 1 Hypothesenbildung
- 2 Versuchsplan
- 3 Versuchsaufbau
- 4 Versuchsdurchführung
- 5 Auswertung (Datenanalyse)
- 6 Schlussfolgerung (Antworten)

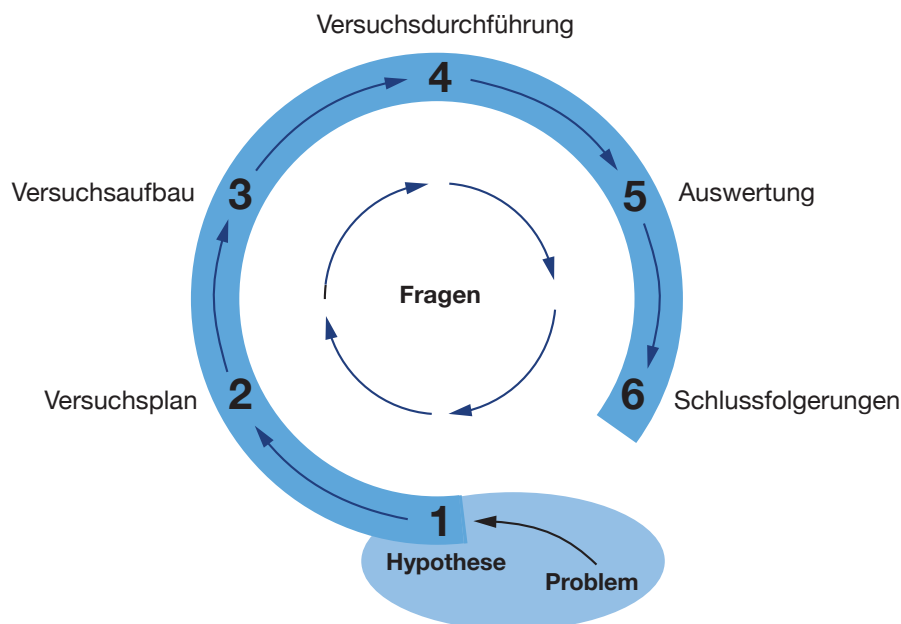


Abbildung 4.1: Wissenschaftliche Methode (Reiß/Sarris, 2012: 60f.)

#### 4.1

### Hypothesenbildung

Nach dem theoretischen Teil, der Problemerkennung, der Bestimmung des Kontexts des Systems, und einer ersten Analyse ergibt sich bereits ein klares Bild. Sowohl aus den theoretischen Grundlagen als auch aus unserer Studienvorbereitung können wir an dieser Stelle unsere Hypothesen in ihrer ersten Form bilden.

Hypothesen werden aus allgemeinen Fragestellungen (Problemen) umgeformt. Die Aussageform beinhaltet einen Lösungsentwurf auf die gestellte Frage und gibt damit eine mögliche Antwort. Die Ergebnisse des darauffolgenden Experimentes entscheiden über die Gültigkeit des Ansatzes. Die Hypothesenbildung erfolgt am Anfang des Design-Experimentes (Reiß/Sarris, 2012: 62f.).

Zu diesem Zeitpunkt sind dies unsere Hypothesen:

«Der SBB Webshop braucht eine Fortschrittsanzeige, um den Benutzer während der Wartezeiten **informiert** zu halten».

«Die Fortschrittsanzeige muss nicht den ganzen Bildschirm blockieren, vielmehr kann sie in dem Bereich **positioniert** werden, in dem die Veränderung stattfindet».

«Die Fortschrittsanzeige kann **kreativer** und ungewöhnlicher gestaltet werden, so dass der Benutzer die Ladedauer als kürzer empfindet».

Diese Hypothesen werden durch neu gewonnene Erkenntnisse (siehe Kapitel 6) und Diskussionen im Team kontinuierlich iteriert.

#### 4.2

### Statistischer Versuchsplan

Der statistische Versuchsplan, der ein wichtiges Dokument für den weiteren Ablauf darstellt, wird nach der Hypothesenbildung aufgestellt, um zu definieren welche Ressourcen wir für unser Experiment einsetzen müssen und, um aufzugliedern, welche Daten wir aus dem Experiment extrahieren wollen und ob sich unsere formulierten Lösungen schliesslich als korrekt erweisen. Dieser Plan ist ein Bindeglied zwischen den vorgängig grob formulierten Hypothesen, dem Versuchsaufbau, der Versuchsdurchführung und der statistischen Auswertung (Reiß/Sarris, 2012: 76f.).

Im Dokument *Versuchsplan* müssen folgende Faktoren schriftlich festgehalten werden (Anhang 9):

- Wie viele Faktoren wollen wir untersuchen? (3 Hypothesen bisher, später um eine vierte ergänzt)
- Wie wollen wir die Faktoren untersuchen? (Quantitative Analyse via Online-Umfrage)
- Welche Zusatzinformationen sind notwendig? (Das Team kann die Teilnehmer nicht nach Merkmalen auswählen und hat keinen Einfluss auf deren Umgebung, während sie den Test machen)
- Welche Art von Experiment führen wir durch? (Quasi-Experiment)

Wir haben ein quasi-experimentelles Design für unser psychologisches Experiment bestimmt. Bei der Bestimmung des Designs für dieses Experiment, müssen folgende Bedingungen erfüllt sein (Reiß/Sarris, 2012: 80f.):

**1** Vor Versuchsbeginn muss eine kausaltheoretische Hypothese vorhanden und begründet sein

Unsere Hypothesen sind bereits vorhanden, werden aber im Laufe des Versuchsaufbaus (siehe Kapitel 6) verfeinert.

**2** Experimentelle Variablen können von uns beeinflusst werden

Folgende Merkmale stellen potentielle experimentelle Variablen für unsere Studie dar:

- Zeit
- Positionierung der Fortschrittsanzeige
- Typ der Fortschrittsanzeige

In unserem Test untersuchen wir eine dieser Variablen: Die vier Prototypen unterscheiden sich durch die Visualisierung der Fortschrittsanzeige. Die Ladedauer und die Positionierung sind bei allen Prototypen gleich. In Kapitel 5.2 beschreiben wir detailliert die Definition der von uns definierten Variable.

**3** Alle übrigen Versuchsbedingungen kontrollierbar bzw. kontrolliert

Leider sind nicht alle Versuchsbedingungen kontrollierbar. In unserem Fall ist es z.B. so, dass wir keinen Einfluss auf die Teilnehmer und ihre Umgebung ausüben können. Damit können wir mögliche Störfaktoren nicht beeinflussen wie es in einer kontrollierten Testumgebung wie einem Test-Labor möglich gewesen wäre. Wir wissen nur wenig über die Merkmale der Testbenutzer. Da die verzögerte Ladedauer bei allen Benutzern auftreten wird, haben wir die folgenden Faktoren nicht berücksichtigt: Geschlecht, Herkunft und Bildungsstand, wie oft die Personen mit dem

Zug fahren oder ob sie ein General Abonnement besitzen.

Weitere mögliche Störfaktoren, wie zum Beispiel der Lärmpegel in der unmittelbaren Umgebung des Benutzers, wenn der Test zum Beispiel im Büro ausgeführt wird, Lichtverhältnisse oder die Geschwindigkeit des Browsers, können wir nicht beeinflussen.

Da alle Bedingungen ausser der Kontrolle der Versuchsbedingungen (alle übrigen Versuchsbedingungen kontrollierbar bzw. kontrolliert) erfüllt sind, folgt unser Experiment einem quasi-experimentellem Design.

### 4.3

## Versuchsaufbau

Nachdem wir den Design-Typ für unsere Studie bestimmten, die Hypothesen formulierten und beides in unserem Versuchsplan festhielten, ist es an der Zeit, alle vorherigen und darauffolgenden Arbeitsschritte aufzulisten.

Wir erstellen eine Checkliste, damit wir mögliche Fehler vermeiden können. Die Liste muss mit den vorherigen Stadien (1 Hypothesenbildung und 2 Versuchsplan) abgestimmt sein. Dasselbe gilt für darauffolgenden Stadien (4 Versuchsdurchführung, 5 Auswertung und 6 Schlussfolgerung). Wir arbeiten iterativ und haben daher schon einige Punkte auf der Checkliste erledigt (Reiß/Sarris, 2012: 102f.).

Die abgebildete Checkliste zeigt in blau, was wir zum Zeitpunkt der Erstellung der Liste bereits ausgeführt haben und in gelb, was als nächstes geplant ist (Abbildung 4.2).

Unter dem Punkt Vorversuch (siehe Checkliste) wenden wir in einer qualitativen Analyse die Methoden Usability-Test, Walkthrough und Hallway Test an. Die qualitative Analyse wird effektiv zu diesem Zeitpunkt eingesetzt, da sie ein offenes und flexibles Vorgehen anbietet, uns erlaubt die Hypothesen weiterzuentwickeln und die Möglichkeit besteht, detaillierteres Feedback von einer limitierten Anzahl von Teilnehmern einholen zu können. Bedenken muss man natürlich, dass die Aussagen der Teilnehmer auf ihren persönlichen Motiven beruhen und dass die Daten nicht repräsentativ für unsere Studie sind. Diese Daten werden deshalb auch nicht statistisch ausgewertet. Damit wir unsere Hypothesen mit einer möglichst grossen Stichprobe testen können, müssen wir ebenfalls auch eine quantitative Analyse durchführen (Abbildung 4.3).

### Quantitatives Vorgehen

### Qualitatives Vorgehen

Messung von empirischen Daten, basiert auf realen Fakten	Relativ offenes und flexibles Vorgehen (oft ist nur ein grober thematischer Leitfaden gegeben)
Untersuchung von grossen Stichproben	limitierte Anzahl von Untersuchungspersonen können teilnehmen aber dafür detailliertere Analyse und Beobachtung
Objektive Messung und Berechnung von Sachverhalten	auf das Verstehen von Sinn (etwa von persönlichen Handlungsmotiven, Absichten) bezogen
Messung zählbarer Merkmale	Verfahren zur Entwicklung neuer Hypothesen (oft auf relativ neuen Forschungsgebieten)
Vorgehensweise zum Testen von Hypothesen	Das Team stellt keinen Anspruch auf Repräsentativität
Das Team stellt einen Anspruch auf Repräsentativität	Daten werden nicht statistisch ausgewertet
Auswertung durch statistische Arbeitsmittel	

Abbildung 4.3: Vergleich quantitatives und qualitatives Vorgehen (Lederer, 2015: 112)

**Status      Checkliste**

✓	Arbeitshypothesen formulieren (erster Draft)
✓	Erste Planung zu unserem Forschungsvorhaben / wichtige Literatur dazu lesen
✓	Eine Rohfassung (Notiz) unseres Forschungsvorhaben
●	Variablen definieren
●	Arbeitsmaterial erstellen: Interview Fragen, Fragebogen und Prototypen entwerfen (erster Draft)
●	Hypothesen ausformulieren
●	Vorversuch: qualitative Analyse (Methoden: Walkthrough, Usability-Test und Hallway Test) anwenden
●	Iteration des Arbeitsmaterials: Fragebogen, Prototypen und Hypothesen
●	Versuchsteilnehmer rekrutieren
●	Hauptversuch: quantitative Analyse (Online-Umfrage)
●	Datenanalyse 1 (Häufigkeitsverteilung grafisch dargestellt und Berechnung der statistischen Kennwerte)
●	Datenanalyse 2 Hypothesen überprüfen nach statistischer Signifikanz
●	Ausarbeitung des Versuchsberichts

Abbildung 4.2: Checkliste unserer experimentellen Studie (Reiß/Sarris, 2012: 103)

#### 4.4

### Versuchsdurchführung

Ein weiterer Teil der Planung ist, sich Gedanken zur Versuchsdurchführung zu machen. Der Versuch des Experiments muss so aufgestellt werden, dass Störfaktoren wenn möglich konstant gehalten werden. In Bezug auf unsere Studie bedeutet das, dass der Testleiter während des geplanten Vorversuchs (Usability-Tests) durch die Testleiter-Testpersonen-Dynamik die Daten möglichst nicht verändert. Eine Veränderung kann je nach Sympathien, die der Testleiter seinen Testpersonen entgegenbringt, entstehen. Eine Reflektion darüber ist am Ende des Tests sinnvoll. Als Abhilfe können zum Beispiel mehrere Testleiter eingesetzt werden. Während des Usability-Tests haben wir die Rollen Beobachter und Testleiter mehrmals innerhalb des Teams getauscht.

Da der Versuch selbst eine Online-Umfrage ist, haben wir weder Kontrolle über die Teilnehmergruppe noch über mögliche Störfaktoren, denen die Teilnehmer ausgesetzt sind.

Ein weiterer Punkt auf den wir achten müssen, ist der Detaillierungsgrad der Versuchsanweisung während des Usability-Tests.

Die Instruktion hat das Ziel, den Versuchsteilnehmer zu informieren und zu motivieren, die Aufgabe zu lösen. Sie informiert ihn über folgende Punkte:

- was der Zweck der Untersuchung ist.
- wie die Untersuchung gegliedert ist.
- was für eine Aufgabe er/sie lösen soll.

Diese Aspekte motivieren die Teilnehmer, am Versuch teilzunehmen:

- Sie können einen Beitrag zur Untersuchung leisten
- Sie können einen gewissen Standpunkt zu diesem Thema einnehmen, während sie an der Umfrage teilnehmen.

Der Teilnehmer kann vollständige, unvollständige und falsche Instruktionen erhalten. Dies hängt von der Fragestellung ab. In unserem Fall liefern wir dem Teilnehmer unvollständige Informationen, da vollumfängliche Instruktionen unser Versuchsziel beeinträchtigen würden. Damit bewahren wir die Untersuchungsnaivität der Teilnehmer, die in Bezug auf das kurze Einblenden der Fortschrittsanzeige und deren Erkennungsfaktor zwingend ist (Reiß/Sarris, 2012: 118f.). Während des Versuchs wird der Teilnehmer über die Versuchsabsichten aufgeklärt und es werden ihm keine Falschinformationen geliefert.

#### 4.5.

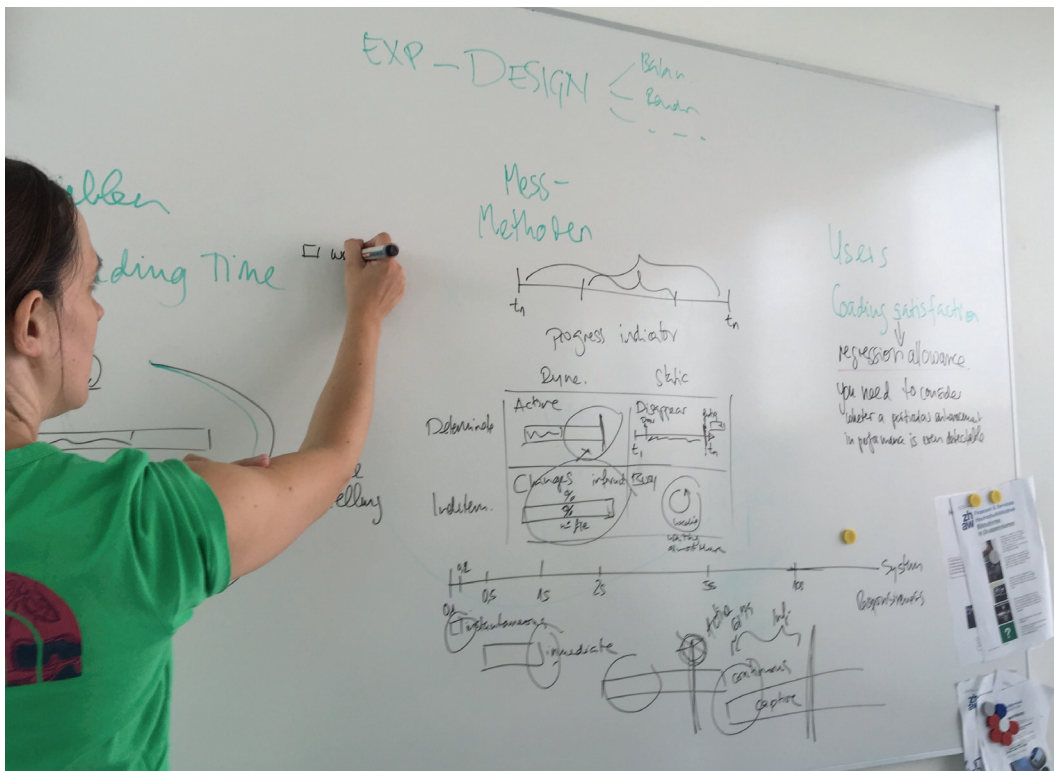
### Auswertung (Datenanalyse)

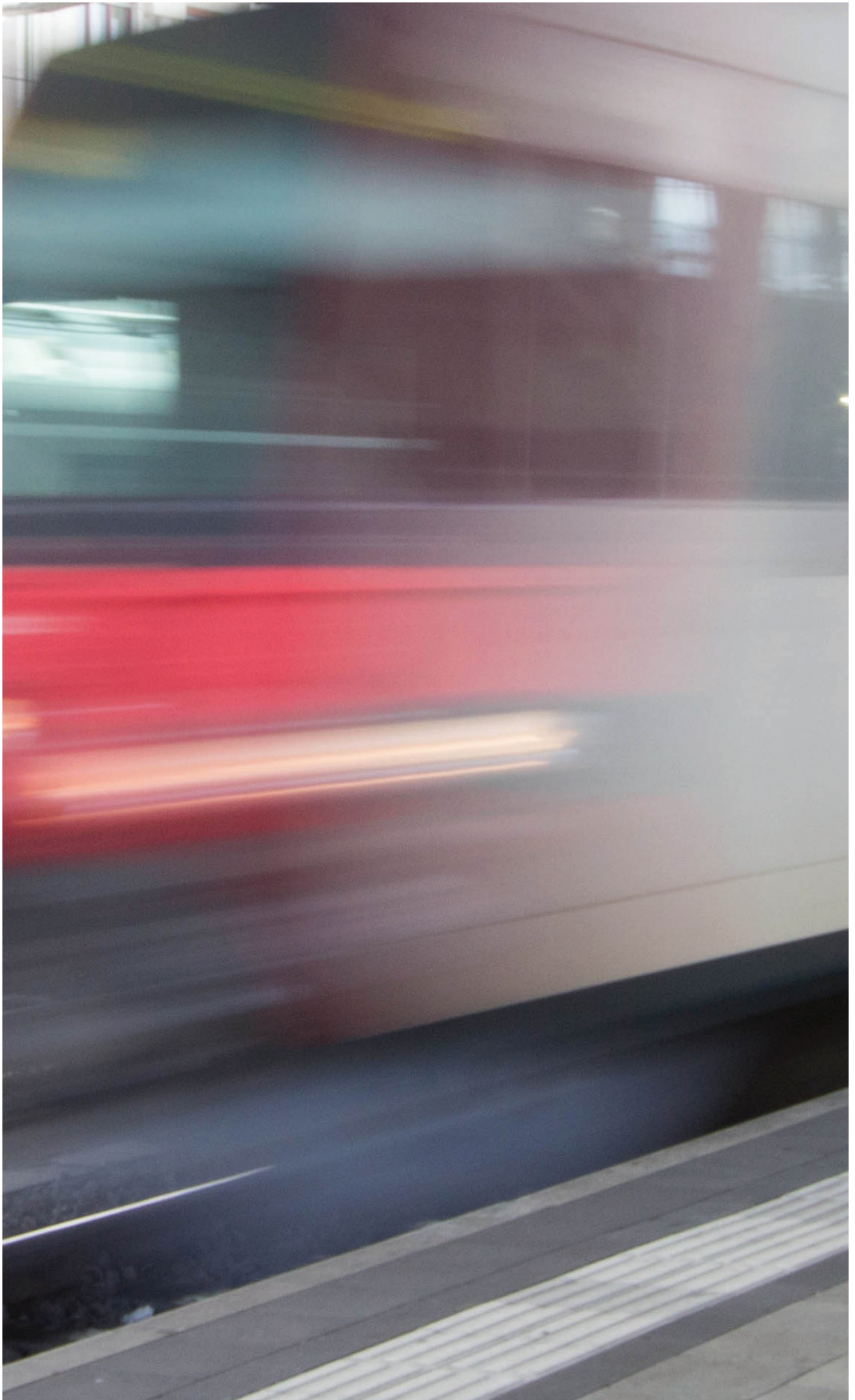
In diesem Stadium werden die gesammelten Daten ausgewertet und analysiert. Im Dokument *Versuchsplan* (siehe 4.2) wurde bereits in einem früheren Stadium festgelegt, was untersucht werden soll. In Kapitel 7 gehen wir ausführlich auf die Auswertung und die Ergebnisse ein und beschreiben, welche Auswertungsmethoden wir anwenden und die Gründe dafür (Reiß/Sarris, 2012: 138f.).

#### 4.6

### Schlussfolgerung (Antworten)

In diesem Stadium werden die Daten interpretiert. Hier finden wir heraus, ob die aufgestellten Hypothesen die erwarteten Antworten liefern. Eine detaillierte Beschreibung folgt in Kapitel 8 (Reiß/Sarris, 2012: 170f.).





# 5. Versuchsaufbau

Nach der detaillierten Planung des Experiments wird in diesem Kapitel der Versuchsaufbau unserer Studie behandelt, die dazugehörige Methodenanwendung im Detail erörtert und die Ableitung des aktuellen Studiendesigns aufgezeigt.

## 5.1

### Definition von User Experience in unserer Arbeit

Bevor wir unsere Hypothesen ausformulieren und unsere Designvarianten ausarbeiten können, definieren wir, was eine optimale User Experience in Bezug auf die Ladezeit und die Fortschrittsanzeige im Rahmen dieser Arbeit bedeutet.

Wenn die Ladezeit angemessen ist, sind die Benutzer zufrieden.

- 1 Je kürzer der Benutzer die Ladezeit empfindet desto zufriedener ist er.

Benutzerzufriedenheit definieren wir mit vier Adjektiven: kurz, angenehm, nützlich und übersichtlich.

- 2 Der Benutzer ist zufrieden, wenn er die Fortschrittsanzeige als angenehm, nützlich und übersichtlich empfindet.

Die optische Attraktivität einer Fortschrittsanzeige definieren wir mit den Adjektiven sympathisch, originell, spannend und attraktiv.

- 3 Der Benutzer findet eine Fortschrittsanzeige optisch attraktiv wenn sie für ihn sympathisch, originell, spannend und attraktiv wirkt.

Eine aussagekräftige Fortschrittsanzeige bedeutet im Rahmen unserer Arbeit für den Benutzer, dass der Ladevorgang für ihn voraussagbar (er weiss, wie lange der Ladevorgang noch andauern wird) und verständlich ist (er versteht, was passiert, wenn er die Fortschrittsanzeige sieht).

## 5.2

### Designvarianten

Im Kapitel 2.4 *Bedeutung des theoretischen Hintergrunds für unsere Arbeit* präsentieren wir die drei Merkmale, die unsere Arbeit beeinflussen: Zeit, Typ der Fortschrittsanzeige und Position der Fortschrittsanzeige. In diesem Kapitel beschreiben wir diese drei Faktoren detailliert und legen unsere Entscheidungen dazu dar.

#### 5.2.1

##### Zeit

Da die Reiseoptionenseite ein neues Konzept beinhaltet und bisher noch keinen realen Bedingungen ausgesetzt wurde, richten wir uns nach den aktuellen Zeiten der Live-Version des Webshops, den die SBB für vergleichbar halten. Die SBB kennen die durchschnittliche Ladedauer zwischen verschiedenen Seiten im aktuellen Billett-Einkaufsprozess (siehe Abbildungen 5.1 und 5.2) und stellen uns diese als Vergleichswerte zur Verfügung.

Wir erhalten von den SBB 2 Messwerte, die für die neue Reiseoptionenseite relevant sind:

- 4.8 Sekunden (durchschnittlich): Nach dem Klick auf *Preis/Kaufen* im Online-Fahrplan.
- 8 Sekunden (durchschnittlich): Nach dem Klick auf *Weiter* auf der Reisende-Seite.

Nach Rücksprache mit den SBB entscheiden wir uns, im Prototypen eine Wartezeit von 7 Sekunden zu verwenden. Mit diesem Wert liegen wir klar in Seows *Captive*-Bereich (>5 Sekunden), auch als *Attention Span* bekannt (Mishunov, 2015a). Wir nehmen somit an, dass 7 Sekunden vom Benutzer als Verzögerung wahrgenommen werden.

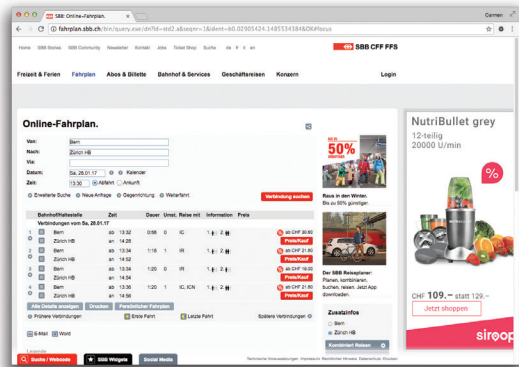
#### 5.2.2

##### Positionierung der Fortschrittsanzeige

In der aktuellen Entwicklungsversion der Website finden wir bereits eine Ladeanzeige zur Überbrückung der Wartezeit. Gemäss den SBB gab es keinen bewussten Entscheid für die Anzeige an dieser Stelle. Trotzdem analysieren wir sie und können sie wie folgt beschreiben (Abbildung 5.3):

- Es handelt sich um eine Anzeige vom Typ *Class B* in Form eines drehenden Rads (Seow, 2008: 80f.)
- Sie erscheint jedesmal, wenn der Warenkorb sich aktualisiert, nachdem der Benutzer eine Änderung in seinen/ihren Reiseoptionen gemacht hat. Sie blockiert dafür die ganze Seite.

Wir empfinden diesen letzten Punkt des Verhaltens der Ladeanzeige als problematisch, vor allem wegen der Auffälligkeit, da das Blockieren des Bildschirms viel Aufmerksamkeit auf die Ladeanzeige und somit auf die Wartezeit lenkt (Zumbrunnen, 2015).



4.8 Sekunden

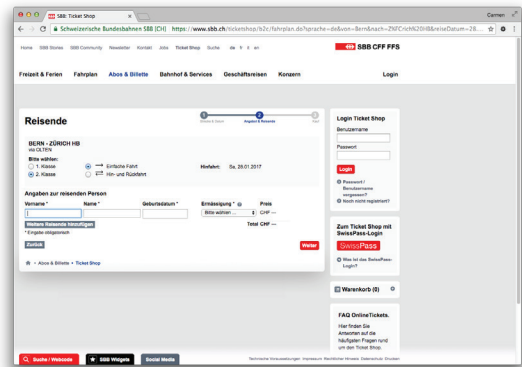
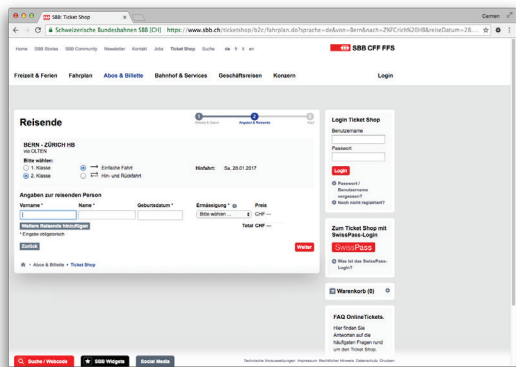


Abbildung 5.1: Durchschnittsladendauer zwischen der Seite Online-Fahrplan und der Seite Reise



8 Sekunden

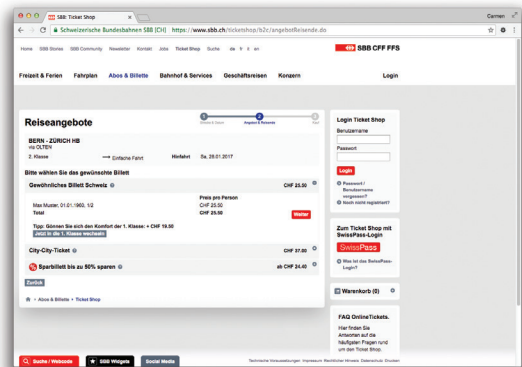


Abbildung 5.2: Durchschnittsladendauer zwischen der Seite Reise und der Seite Reiseangebote

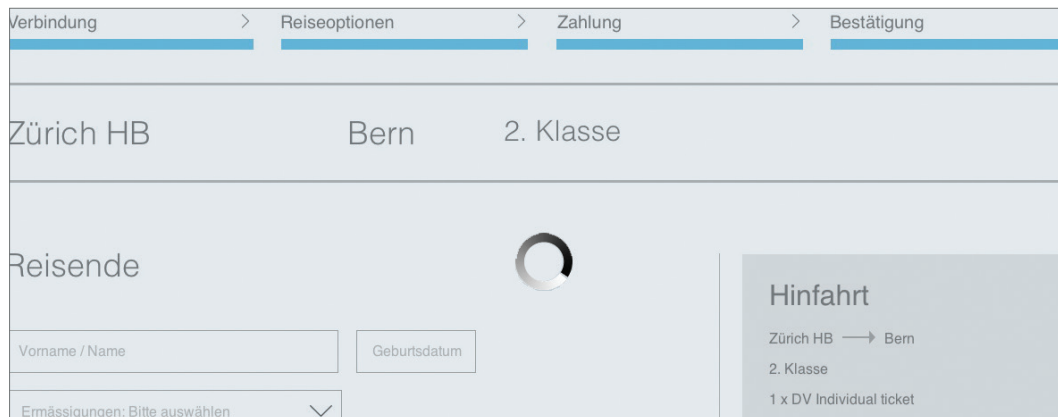


Abbildung 5.3: Aktuelle Fortschrittsanzeige der SBB

Unsere erste daraus folgende Entscheidung ist, verschiedene Positionierungen zu testen. Wir bestimmen drei Positionen auf der Website, wo die Ladeanzeige platziert werden kann:

- 1 Auf der ganzen Seite.
- 2 Auf dem Warenkorb, wo die Änderung stattfindet.
- 3 Auf dem Feld, das die Änderung verursacht.

Die Position 1 ist der Status quo der aktuellen Entwicklungsumgebung, wie oben beschrieben. Die Position 2 und 3 lassen sich einerseits aus unseren im Kapitel 3.2.2 beschriebenen Beobachtungen der Konkurrenz, andererseits aus Van Heals Buch ableiten: «Animating feedback can open up opportunities to provide feedback on interactions as they are happening without disrupting the entire interface visually» (Heal, 2016: 116).

### 5.2.3 Typ der Fortschrittsanzeige

Im Kapitel 2.2.1 beschrieben wir die verschiedenen Typen von Fortschrittsanzeigen, die in der Literatur vorkommen. Seows Klassifizierung (2008: 80f.) und die Erweiterung von Denys Mishunov aus dem Smashing Magazine (2015c) gelten als Grundlage für unsere Entscheidungen für diese Arbeit.

Die Fortschrittsanzeigen, die wir in unserer Studie verwenden, sind folgende (Abbildung 5.4):

- Fortschrittsbalken oder *Class A* mit Prozentanzeige (Seow, 2008: 83).
- Dynamische *Class C*, gebaut aus einer *Class B*-Anzeige mit dynamischem Text, die dem Benutzer Informationen über den Prozess gibt. Das ist für unsere 7-Sekunden-Wartezeit die ideale Anzeige gemäss der Literatur (Mishunov, 2015c).
- *Busy*-Anzeige (*Class B*, Seow, 2008: 83) in Form eines Zuges, der sich von links nach rechts wiederholend bewegt, als spielerische Variante von einem Spinner. Hier wird zusätzlich die Theorie der visuellen Ästhetik nach Norman überprüft, dass attraktive Dinge besser funktionieren. Attraktivität und Ästhetik können nicht endgültig festgelegt und definiert werden. Daher beschränken wir uns bei der Gestaltung auf eine möglichst minimalistische Form, um Störfaktoren durch Effekte wie Farben und Dekorationen weitestgehend auszuschliessen.

Nach Rücksprache mit unserer Auftraggeberin fügen wir eine weitere Variante ein, welche keine Fortschrittsanzeige hat, aber trotzdem wie in den anderen Varianten auch 7 Sekunden Wartezeit hat.

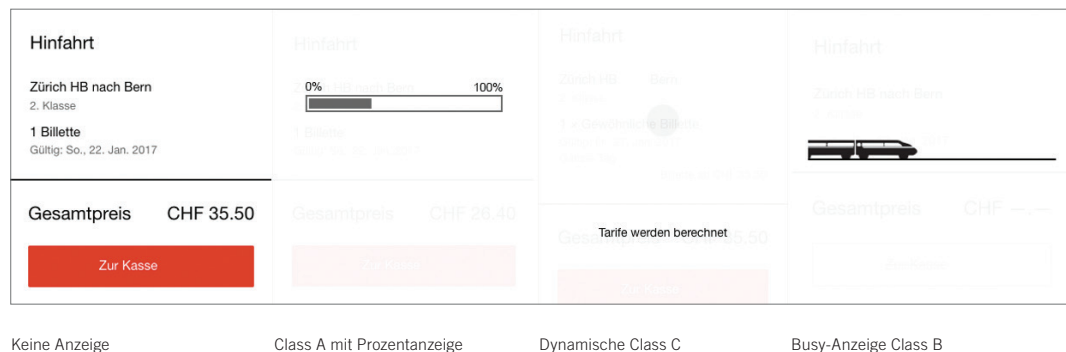


Abbildung 5.4: Fortschrittsanzeigen der Studie

#### 5.2.4.

### Unsere Variable für die Studie

Wie in den vorherigen Kapiteln erklärt, identifizieren wir drei wichtige Punkte, die einen Einfluss auf unsere Studie haben.

Da wir mit einer fixen Zeit von 7 Sekunden arbeiten, bleiben uns die beiden Merkmale Typ und Position, die uns die Möglichkeit bieten, zwischen mehreren Optionen zu variieren.

Wir realisieren schnell, dass wir nicht mit zwei Variablen arbeiten können. Damit wir alle möglichen Varianten testen können, würden wir eine sehr grosse Stichprobe benötigen, was für unsere Ressourcen nicht realistisch ist. Die Bearbeitung des Datensatzes würde auch komplexer werden und wir entscheiden uns für ein einfacheres Experiment mit nur einer Variablen. Deswegen wird die Position der Fortschrittsanzeige fixiert.

Die Positionierung der Fortschrittsanzeige auf der ganzen Seite (den Status quo) empfinden wir als ungünstige Lösung, da sie zu viel Aufmerksamkeit auf die Anzeige lenkt. Das ist kontraproduktiv für die wahrgenommene Wartezeit (Branaghan/Sanchez, 2009: 528f., Zumbrennen, 2015). Somit ist diese Option eliminiert.

Die Position 3 (neben dem Feld) hat als Nachteil, dass sie nicht zu allen unseren Designs für die Fortschrittsanzeigen passt. Aus praktischen Gründen eliminieren wir auch diese Position. Wir finden, dass die Position 2 die am besten passende ist, und somit wird für die Studie die Position der Ladezeitanzeige auf dem Warenkorb fixiert (Abbildung 5.5).

Basierend auf der Variablen *Typ der Fortschrittsanzeige* bilden wir unsere Hypothesen und entwickeln entsprechend die Prototypen und Fragebogen für die Studie.

Typ / Position	Ganze Seite	Warenkorb	Feld
Class A (Fortschrittsbalken)	–	✓	–
Class C (Busy-Indikator mit Feedback-Text)	–	✓	–
Class B (spielerische Variante)	–	✓	–
Keine Anzeige	–	✓	–

Abbildung 5.5: Variable der Studie

### 5.3

## Hypothesenformulierung

Abgeleitet vom oben beschriebenen gesammelten literarischen Wissen, von Erkenntnissen erhalten durch Stakeholder Interview und der Konkurrenzanalyse werden die bereits gebildeten Hypothesen zu diesem Zeitpunkt ausformuliert und ergänzt. Diese Hypothesen werden zuerst mit einer qualitativen Analyse (Usability-Test) und danach mit einer quantitativen Analyse (Online-Umfrage) überprüft.

Aufgrund unserer Theoriebildung ist unsere wichtigste Hypothese, dass während des Ladevorgangs von sieben Sekunden die optimale Darstellung ein Spinner mit einer textuellen Rückmeldung ab vier Sekunden ist (Prototyp 3). Wir nehmen an, dass der Benutzer mit dieser Darstellung am zufriedensten ist und dass er die Dauer als angemessen empfindet (siehe Kapitel 5.2.3). Eine andere Hypothese bezieht sich auf die visuelle Ästhetik. Hier wollen wir herausfinden, welche Fortschrittsanzeige dem Benutzer visuell am besten gefällt (wahrscheinlich Prototyp 4). Zum Schluss interessiert uns, ob dem Benutzer eine Fortschrittsanzeige überhaupt hilft, die Zeit zu überbrücken und ob er die Zeit dadurch kürzer einschätzt.

Um diese Hypothesen durch eine Online-Umfrage zu prüfen, bauen wir vier Prototypen, die sich nur durch die Fortschrittsanzeige unterscheiden (Abbildung 5.6, siehe Kapitel 6.1). Für uns ergeben sich die folgenden Hypothesen:

- 1 Eine Ladeanzeige hilft, dass der Benutzer die Ladezeit als kürzer einschätzt, bzw. wahrnimmt.

Das Ziel unserer Masterarbeit ist, herauszufinden, wie man durch die Gestaltung der Fortschrittsanzeige die wahrgenommene Dauer verkürzen kann (Branaghan/Sanchez 2009: 537, siehe Kapitel 2).

Diese Hypothese stützt sich auf die Literaturrecherche über Responsiveness und User Experience, die empfiehlt, den Benutzer immer zu unterhalten und informiert zu halten. Gemäss Myers (1985: 17) wünschen Benutzer sich Fortschrittsanzeigen. Die Art, wie eine Fortschrittsanzeige sich präsentiert und bewegt, kann die Zeitwahrnehmung der Benutzer beeinflussen (Harrison et al., 2007: 115). Wie bereits erwähnt, nehmen wir an, dass der Benutzer eine Version mit Fortschrittsanzeige einer ohne Fortschrittsanzeige bevorzugt. Deswegen haben wir auch einen Prototypen ohne Fortschrittsanzeige gebaut (siehe Kapitel 7.1), um den Unterschied zu den Prototypen 2, 3 und 4 mit Fortschrittsanzeige zu messen.

**2** Die Zufriedenheit mit der visuellen Ästhetik der Ladeanzeige ist beim Prototypen 4 am grössten.

Die Idee entstand während der Definition des Kontexts und der Konkurrenzanalyse (siehe Kapitel 3.2). Die Swiss zeigt während des Ladevorgangs Flugzeuge. Wir dachten, dass ein kontinuierlich fahrender Zug als Fortschrittsanzeige zu den SBB passt. Während des Ladevorgangs in Prototyp 4 fährt ein Zug in einer spielerischen Art über den Warenkorb (Abbildung 5.7). Der Prototyp 4 hebt sich damit von den anderen Prototypen ab.

Unser Empfinden, dass der visuelle Faktor zur Zufriedenheit des Benutzers beitragen wird, finden wir durch die Literatur unterstützt. Gemäss Donald Normans These «Attractive things work better» (Norman, 2004: 17) reagieren Benutzer auf eine visuell attraktive Anzeige besser als auf eine Anzeige, die gewöhnlicher scheint. Im Weiteren beschreibt Norman die drei Ebenen der Wahrnehmung und wie der Mensch auf eine positive und negative Art bewusst und unterbewusst Dinge wahrnimmt (Norman, 2004: 3) (siehe Kapitel 3.2.2). Zwei andere Autoren, Thielsch und Jaron, beschreiben, wie der Benutzer drei zentrale Dimensionen wahrnimmt wenn er eine Website beurteilt: Inhalt, Usability und Ästhetik (Thielsch/Jaron, 2008: 129f., Thielsch, 2008a: 60f.).

**3** Die Benutzerzufriedenheit ist beim Prototypen 3 am grössten.

Wir definieren Benutzerzufriedenheit ähnlich der Loading satisfaction von Lallemand (Lallemand/Gronier, 2012). Hier wurden die geschätzte Dauer der Ladezeit, die Angemessenheit der Ladeanzeige, die Zufriedenheit mit der Ladeanzeige und der Fokus auf die Ladeanzeige als entscheidende Faktoren definiert.

Im vorliegenden Fall wird die Zufriedenheit mit den folgenden Skalen definiert:

#### **Wahrgenommene Dauer**

Die wahrgenommene Dauer der Ladezeit ist ein wichtiger Bestandteil der Messung. Sie erscheint uns entscheidend, um die Zufriedenheit und den Erfolg der Fortschrittsanzeige zu messen.

#### **Nützlich/Nutzlos**

Hier soll getestet werden, ob die Ladeanzeige dem Benutzer sinnvoll erscheint und eine nützliche Rückmeldung bietet. Somit soll im Sinne des ISO-9241-110-Standards die positive Rückmeldung des Systemzustandes evaluiert werden.

### **Übersichtlich/Verwirrend**

Die Ladeanzeige soll den Benutzer nicht überfordern, ablenken oder ihn anderweitig aus seinem Gedankenstrom reißen. Dies ist ein Kernelement einer erfolgreichen Interaktion (Doherty, 2015).

### **Angenehm/Lästig**

Die Zufriedenheit mit der Ladeanzeige geht über Usability-Aspekte hinaus. Im Sinne der Definition von User Experience in dieser Arbeit sollen hier emotionale Faktoren gegenüber der Ladeanzeige mitbewertet werden.

Unsere Hypothese lautet, dass die Fortschrittsanzeige des Prototyps 3 die Zufriedenheit mit dem Ladeverhalten erhöht. Hintergrund ist die Benutzerzufriedenheit in Bezug auf die Ladedauer. Bei Prototyp 3 erhalten Benutzer Feedback sowohl visuell als auch in der Form von relevanter textueller Beschreibung (Mishunov, 2015c).

Der Benutzer ist zufrieden, wenn er die Ladezeit als kürzer und die Fortschrittsanzeige als angenehm, nützlich und übersichtlich (siehe Kapitel 5.1) empfindet.

**4** Die Aussagekraft der Ladeanzeige ist bei Prototyp 3 am stärksten (bei Prototyp 1 am schwächsten).

Der Benutzer braucht eine aussagekräftige Ladeanzeige und die Rückmeldung trägt zum allgemeinen Verständnis der Benutzung der Applikation bei (Branaghan/Sanchez 2009:528f).

Nach unserer Definition (siehe Kapitel 5.1) ist die Aussagekraft der Fortschrittsanzeige für den Benutzer am stärksten, wenn er eine Annahme treffen kann, wie lange der Ladevorgang noch dauert, ihm verständlich ist, was genau im Hintergrund passiert, zum Beispiel welche Daten geladen werden, und dass er den Ladevorgang als schnell empfindet. Wenn er die Fortschrittsanzeige als schnell empfindet, kommt ihm auch die Ladezeit kurz vor.

<b>Die vier Prototypen:</b>	<b>Beschreibung</b>
Prototyp 1	Ohne Fortschrittsanzeige.
Prototyp 2	Mit Fortschrittsanzeige: Ladebalken.
Prototyp 3	Mit Fortschrittsanzeige: Spinner und textuelle Rückmeldung ab vier Sekunden.
Prototyp 4	Mit Fortschrittsanzeige: Bildliche Darstellung eines Zugs als Fortschrittsanzeige.

**Abbildung 5.6:** Unsere Prototypen



**Abbildung 5.7:** Fortschrittsanzeige mit dem Zug



# 6. Durchführung der Studie

In in diesem Teil besprechen wir das Design der vier unterschiedlichen Prototypen, die Iteration der Lösungsansätze und die Durchführung unserer Studie.

In dieser Phase werden mit den bisher erlangten Informationen zuerst Konzepte und dann Design-Lösungen konzipiert. Das ist ein iterativer Prozess, der verschiedene Runden durchläuft, bis ein zufriedenstellender Entwurf vorliegt. Artefakte, die aus diesem Prozess entstehen, können Designdokumente, Papier-Prototypen (low-fidelity Prototypen) oder interaktive Prototypen (high-fidelity Prototypen) sein.

## 6.1 Prototyp

Nach der Hypothesenbildung und der Operationalisierung (siehe Kapitel 4) geht es nun an die Erstellung des Prototypen.

Mit einem interaktiven Prototypen kann man auf einfache und effiziente Weise eine Interaktion herstellen und sie mit dem Benutzer evaluieren. Der Prototyp demonstriert, wie der Benutzer auf der Desktop-Seite agieren kann. Durch wiederholte Iterationen wird die Design-Lösung optimiert. Es müssen nicht alle Interaktionen dargestellt werden. Besonders wichtig ist es, den Prototypen an den Stellen klickbar zu machen, an denen wir Fragen haben und Benutzer-Feedback abholen möchten. Je nach Projektaufwand und Zeit können unterschiedliche Tools verwendet werden.

Axure ist ein weit verbreitetes Programm, das wenig Programmiererfahrung voraussetzt und leicht einsetzbar ist. Der Nachteil von Axure-Prototypen ist die Zeitverzögerung, die je nach Browser

anfällt und die in unserem Fall auf Sekunden genau sein muss. Demgegenüber stehen HTML-Prototypen, welche mehr Kontrolle bieten, Zeitverzögerungen zu eliminieren (Goodwin, 2009: .617f.).

Da die Ladezeit den wichtigsten Faktor darstellt und präzise kontrolliert werden soll/muss, entwickeln wir für unsere Lösung vier detaillierte interaktive HTML-Prototypen. Die Prototypen sind alle gleich aufgebaut und unterscheiden sich nur durch eine Variable (die Fortschrittsanzeige) (siehe Kapitel 6.1).

### 6.1.1 Ausarbeitungsgrad

Aufgrund der Hypothesen sind die Dimensionen der Prototypen relativ schnell klar. Zum einen sollte der visuelle Ausarbeitungsgrad nicht zu hoch sein, um die Benutzer nicht mit visuellen Details abzulenken und zum anderen sollten sich die Erstellung der Prototypen sowie gegebenenfalls nötige Anpassungen nicht zu aufwändig gestalten.

### 6.1.2 Technische Anforderungen

Im Hinblick auf die Funktionalität gelten für die Prototypen spezielle technische Anforderungen. Einerseits ist zwar die funktionale Tiefe nicht sehr gross, da keinerlei Datenbankverbindung oder ähnliches getestet werden muss. Andererseits ist aber der Anspruch an die Frontend-Technik recht hoch, da die Rahmenbedingungen, insbesondere die Ladedauer für jeden Benutzer an seinem Endgerät möglichst genau kontrolliert und festgelegt werden sollen.

Somit entschieden wir uns gegen die Benutzung eines Prototypen-Tools, mit der zwar die Erstellung des Prototypen an sich vielleicht schneller gegangen wäre, die Technik aber schwerer zu kontrollieren ist. Ein weiterer Pluspunkt für eine HTML-Variante der Prototypen ist der Zugang zu dem vorhanden SBB-Ticketshop-Prototypen, aus dem wir bereits vorhandene HTML-Elemente übernehmen können.

Zum Einsatz kam ein modernes Javascript-Framework (Aurelia.js), so dass wir unsere Prototypen als Single-Page-App (SPA) anbieten können. SPAs bieten den Vorteil, dass nur beim erstmaligen Laden der Seite eine Netzwerkverbindung notwendig ist, mit der der gesamte Prototyp auf einmal heruntergeladen werden kann. Somit lässt sich unabhängig von der Netzwerkverbindung des Benutzers die Länge der Ladedauer und das Interaktionsverhalten genauestens emulieren und es lassen sich jederzeit ähnliche Nutzungsbedingungen sicherstellen, was die Störfaktoren schlechte oder unterbrochene Internetverbindung ausschliesst.

Gleichzeitig wird durch eine Stylesheet-Beschränkung per CSS sichergestellt, dass Benutzer nur ab einer bestimmten Mindestbreite des Bildschirms Zugriff auf den Prototypen bekommen. Dadurch wird das Experiment auf grosse Bildschirme beschränkt, wodurch davon auszugehen ist, dass so gut wie alle Teilnehmer das Experiment an einem Laptop oder Desktop-PC

durchführen, was unseren Anforderungen an das Experiment entspricht. Dabei wird das responsive Verhalten innerhalb dieser Bildschirmgrößen, das heisst die automatische Anpassung des Webseiten-Layouts an Bildschirmgrößen vom kleineren bis zum grossen Laptop oder Desktop-PCs, weitestgehend dem vorhandenen SBB-Prototypen nachempfunden, um den Benutzungs-Kontext möglichst ähnlich zum SBB-Prototypen zu halten und relevante Erkenntnisse erzielen zu können.

Zudem haben wir durch diese Technik die Möglichkeit, den Prototypen auch ohne bestehende Netzwerkverbindung, z.B. bei Usability-Tests, verwenden zu können.

### 6.1.3 Benutzerführung im Remote-Testing

Die Heranführung der Testnutzer an die Benutzung der Prototypen stellt eine besondere Herausforderung für uns dar, da unser Plan vorsieht, die Benutzer per unbeobachtetem Remote-Test das Experiment von ihrem eigenen Gerät aus durchführen zu lassen. Somit ist eine möglichst effektive wie effiziente Einführung in das Testzenario und die Aufgabenstellung nötig.

1		Zürich HB	ab 17:08	1:18	1	IR	1.	2.	ab CHF 25.00
		Bern	an 18:26						<b>Preis/Kauf</b>
2		Zürich HB	ab 17:32	0:56	0	IC	1.	2.	ab CHF 25.00
		Bern	an 18:28						<b>Preis/Kauf</b>
3		Zürich HB	ab 17:38	1:18	2	RE, ICE	1.	2.	ab CHF 25.00
		Bern	an 18:56						<b>Preis/Kauf</b>
4		Zürich HB	ab 17:55	1:26	0	IR	1.	2.	ab CHF 25.00
		Bern	an 19:21						<b>Preis/Kauf</b>
5		Zürich HB	ab 18:02	0:56	0	IC	1.	2.	ab CHF 25.00
		Bern	an 18:58						<b>Preis/Kauf</b>

Alle Details anzeigen
Drucken
Persönlicher Fahrplan

← Frühere Verbindungen
Erste Fahrt
Letzte Fahrt
Spätere Verbindungen →

Abbildung 6.1: Fahrplanansicht aus dem alten Prototyp

Bei unserer ersten Variante der Prototypen gestaltete sich diese Einführung interaktiv. Der Benutzer war dazu angehalten, selbst Start und Ziel einzugeben und danach eine Verbindung herauszusuchen.

Wie sich aber bei unserem Usability-Test mit dem Prototypen zeigte, gab es zwei Probleme mit dieser interaktiven Einführung. Einerseits waren nur einzelne Eingabefelder, nicht der ganze Nutzungskontext der SBB-Seite auf unseren Prototypen vorhanden, was zu einer ungewohnten und künstlichen Benutzungsumgebung für die Testnutzer führte.

Das zweite Problem stellt die Fahrplanansicht dar (Abbildung 6.1). Hier wurde zuerst ein Screenshot des aktuellen SBB-Ticketshops angezeigt, der durch viele Buttons weitere Interaktionsmöglichkeiten suggerierte, obwohl diese nicht vorhanden waren. Bei jeglichem Klick auf diesen Screenshot wurde zum nächsten Schritt gesprungen, was oftmals für grosse Verwirrung sorgte.

Daher ersetzen wir nach dem Usability-Test die interaktive Einführung mit einem kurzen Szenario, das statt realer Eingaben durch Illustrationen und kurzen Texten die Vorstellungskraft der Testnutzer anspricht, was sehr viel besser ankommt (siehe Kapitel 6.2, Hallway Testing). Screenshots von Version 1 und 2.1 der Prototypen können im Anhang 10 angesehen werden.

## 6.2

### Vorversuch: Usability-Test

Einen Usability-Test vor der Umfrage durchzuführen – als sogenannten Vorversuch – ermöglicht uns, Feedback zu den Prototypen und dem Fragebogen zu erhalten. Die Anwendung dieser qualitativen Methode hat im Vergleich zur Umfrage (quantitativen Methode) den Vorteil, dass wir die Benutzer direkt nach ihrer Meinung fragen können (siehe Kapitel 4.3, Abbildung 4.3).

Ein Usability-Test führt eine Anzahl von Benutzern durch möglichst realistische Szenarios. Der Lösungspfad ist ausschlaggebend, da manche Szenarios zeitliche Abhängigkeiten aufweisen und der Benutzer möglichst den optimalen Weg wählen soll. Das Wichtigste ist, dass der Benutzer während des Usability-Tests laut denkt und der Projektgruppe dadurch möglichst viel Input liefert. Diese Methode kann sehr effektiv sein, da häufig kleinere, aber auch schwerwiegende Probleme frühzeitig identifiziert werden können. Der Usability-Test hat auch den Vorteil, dass wir die Benutzer beobachten und die Reaktionen an ihren Gesichtern ablesen können.

An dieser Stelle wollen wir in Erinnerung rufen, dass eine Evaluierung nie eine perfekte Aussage machen kann, da eine Laborsituation nicht ganz der Realität entspricht. Je intensiver man aber ein Produkt mit verschiedenen Techniken evaluiert, desto näher kommt das Design-Team an eine optimale Design-Lösung.

#### Planung

Einen Usability-Test erfolgreich durchzuführen, erfordert sorgfältige Planung. Festzulegen, welche Fragen das Team beantwortet sehen möchte, ist einer der ersten Schritte. Im weiteren Verlauf wird der Detailgrad des Prototypen festgelegt,

das heisst, ob der Prototyp als low-fidelity oder high-fidelity erarbeitet wird. Wir arbeiten mit einem high-fidelity Prototypen (siehe Kapitel 6.1)

### **Testszzenarien**

Das via Cognitive Walkthrough geprüfte und ergänzte Szenario (siehe Kapitel 6.2.1, Anhang 5) wird für den Usability-Test ausgedruckt, aber auch auf dem Screen angezeigt. Der Benutzer kann nachlesen, was er zu tun hat. Wir wollen die Informationen so detailliert wie möglich halten. Die Definition von Testszzenarien ist ein wesentlicher Aspekt bei Usability-Tests. Sie legen fest, welcher Aspekt getestet werden soll und welche Fragen das Team beantwortet sehen möchte.

### **Rekrutierung der Teilnehmer**

Die Rekrutierung der Teilnehmer kann auf unzählige Arten erfolgen. Man kann sich bei höherem Budget z.B. auf eine professionelle Rekrutierungsagentur verlassen. Wichtig ist in jedem Fall, dass die möglichen Benutzer das Interface vorher nicht gesehen haben und sich damit keine vorherige Meinung bilden können. Die Merkmale der Teilnehmer müssen auf die vorher festgelegten Benutzergruppen abgestimmt sein. Eine Stichprobe von 4-5 Personen kann für eine in kleinem Rahmen definierte Aussage genügen. Das Testen von komplexeren Systemen, wie zum Beispiel einer ganzen Applikation, benötigt eine grössere Gruppe von zum Beispiel 15-20 Personen. In unserem Fall haben wir entschieden, einen Usability-Test mit vier Testpersonen durchzuführen.

### **Usability Räumlichkeiten**

Für den Usability-Test können wir die Räumlichkeiten der UBS verwenden. Das Usability-Lab besteht aus zwei Sitzungszimmern.

Da es sich um einen kleinen Test handelt und die Infrastruktur uns nicht erlaubt, private Rechner an das UBS-Netz anzuschliessen und somit eine Monitorübertragung vom einen zum anderen Raum verhindert wird, legen wir den Beobachter und Test-Raum zusammen. Die beiden Beobachter können vom Hintergrund aus zuschauen und das Gesehene notieren.

### 6.2.1

## Vorbereitung: Cognitive Walkthrough

Der Cognitive Walkthrough (Englisch für *Durchdenken eines Problems*) wird dann angewendet, wenn ein Usability-Experte vor dem eigentlichen Usability-Test eine Usability-Inspektion anwenden möchte. Der Usability-Experte versetzt sich in die Rolle des Benutzers und durchläuft die vorgegebenen Handlungsabläufe. Er nimmt an, dass der zukünftige Test-Benutzer den Weg des geringsten Widerstandes nehmen wird. Der Input des Usability-Experten kann zu einem erfolgreicherem Testverfahren beitragen (Wharton et al., 1994: 105f.).

Unsere SBB-Auftraggeberin schlägt uns die Methode Cognitive Walkthrough vor, da sie sich zur Vorbereitung der Usability-Tests gut anbietet, um mögliche Denkfehler zu vermeiden und mögliche Probleme im Ablauf frühzeitig zu erkennen.

Sie bieten uns dazu Unterstützung von einer Usability-Expertin aus ihrem Team an. In den Räumlichkeiten der UBS, dem Usability-Lab, in dem der eigentliche Usability-Test stattfinden wird, führen wir die SBB-Mitarbeiterin durch den gesamten Testablauf.

Wir wollen folgende Punkte mit ihr anschauen: (Wharton et al., 1994: 105f.)

### Wer sind die Benutzer?

Die Benutzer für den Usability-Test besitzen als gemeinsames Merkmal alle kein GA, benutzen aber den Zug. Dieses Merkmal hat uns der UX-Lead der SBB mitgeteilt.

### Was für ein Szenario soll evaluiert werden?

Das Szenario für den Usability-Test haben wir auf

einem separaten Blatt für die Testbenutzer ausgedruckt (siehe Anhang 5). Es beschreibt in ein paar kurzen Sätzen die Motivation des Benutzers, ein Bahnbillet zu kaufen. Das Ziel ist es, den Benutzer so schnell wie möglich auf die Shop-Seite zu leiten, da dieser Screen die Fortschrittsanzeige zeigt.

### Wie soll der Benutzer das Szenario durchlaufen?

Der Benutzer soll das Szenario lesen und sich danach durch jeden der vier Prototypen klicken.

### Wie ist das Interface definiert? (Was muss der Benutzer sehen, damit er/sie den Weg durchlaufen kann?)

Wir entscheiden uns, dass der Benutzer zwei Eingabefelder hat, um seine Fahrt zu definieren. Der Screen mit den verschiedenen Zugoptionen, der danach angezeigt wird, ist durch ein statisches Bild definiert. Der Benutzer muss sich also nicht für ein Angebot entscheiden und kann irgendwo auf dem Screen klicken, um auf die Reiseoptionenseite weitergeleitet zu werden. Auf der Reiseoptionen-Seite wollen wir sehen, ob er sich gut zurechtfindet und die im Szenario beschriebenen Tickets kaufen kann.

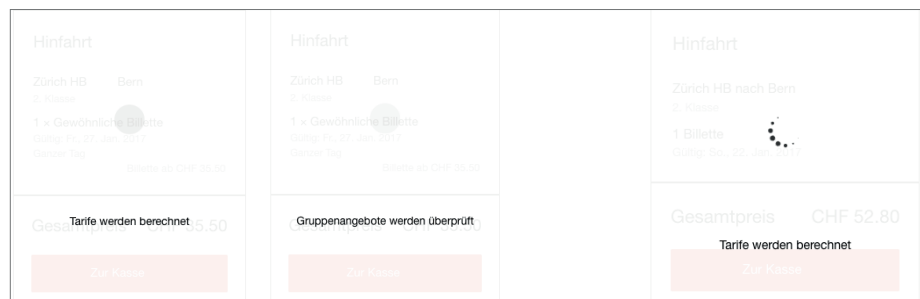
Der Input der Usability-Expertin kann zwei Punkte vor dem Usability-Test aufklären:

- 1 Was muss der Benutzer wissen, damit er ein Szenario ausführen kann?**
- 2 Was soll der Benutzer lernen, während er das Szenario ausführt?**

## 1 Was muss der Benutzer wissen, damit er ein Szenario ausführen kann?

Sie bestätigt uns, dass das Szenario klar beschrieben ist und der Benutzer genug Hintergrundinformationen hat, um das Szenario auszuführen.

- Ein paar hilfreiche Detailbeschreibungen ergänzen das Szenario
  - welche Ermässigungen für die definierten Reisenden zutreffen, zum Beispiel die Junior-Karte für den Enkel und das Halbtax für die beiden Erwachsenen.
  - die Anzahl Billette, die der Benutzer kaufen soll, sollte im Szenario angegeben werden, damit der Benutzer dies nicht selber ausrechnen muss.
- Auf jedem darauf folgenden Screen kann auf Knopfdruck das Szenario ein- und ausgeblendet werden. Sie würde es begrüßen, die Informationen immer sehen zu dürfen.
- Das Szenario sollte auf der linken Seite statt auf der rechten eingeblendet werden.
- Die Willkommensseite soll auf einem separaten Screen angezeigt werden.
- Die Eingabefelder *von* und *nach* sollen leer bleiben, damit der Benutzer die Orte selber eingeben kann und das Szenario realistischer wirkt.
- Visuelle Details wie die Alignierung der Icons lenken den Benutzer vom Test nicht ab.
- Technische Ergänzungen sind:
  - der Kaufknopf soll aktiviert werden, wenn ein Preis angezeigt wird.
  - nach jeder hinzugefügten Person soll der Preis aktualisiert werden.
  - die Eingabe sollte nach Auswahl der Ermässigung blockiert werden. Das heisst, man kann nichts eingeben, bis die 7 Sekunden der Fortschrittsanzeige abgelaufen sind.
- Textliche Änderungen sind:
  - das Datum der Billettbestellung sollte immer auf den aktuellen Tag gesetzt werden.
  - der Text «Bitte achten Sie auf die automatische Preisveränderung im Warenkorb» soll nun lauten: «Am Ende klicken Sie auf "Zur Kasse". Das wird Sie zum anschliessenden Fragebogen führen».



alter Prototyp 3

neuer Prototyp 3

Abbildung 6.2: Alter und neuer Spinner

## 2 Was soll der Benutzer lernen, während er das Szenario ausführt?

Der Benutzer soll mit dem neuen SBB-Konzept vertraut gemacht werden. Sie kann alle Billette bestellen und drückt zum Schluss intuitiv auf den *Kaufen*-Knopf. Die Anzeige, die danach erscheint, («Diese Funktion ist nicht implementiert») verwirrt sie. Der Knopf sollte Ihrer Meinung nach den Übergang zum Fragebogen auslösen.

- Der *Beenden*-Knopf kann dafür gelöscht werden
- Die visuelle Repräsentation des Spinners, der den Text anzeigt, sollten wir ändern zu einem Spinner, der bekannter ist (Abbildung 6.2).
- Wir sollten eine Wartezeit einfügen, wenn die Benutzer ihre Reisedaten eingeben.
- Wenn der Ladebalken als Fortschrittsanzeige erscheint (Abbildung 6.3), sind zu den angezeigten Prozentangaben zusätzliche die effektiven Zahlen einzufügen.

Unter anderem rät sie uns, ein Skript zu benutzen, so dass wir den gleichen Ablauf bei jeder Person befolgen. Basierend auf obigem Feedback passen wir die Prototypen an.

### 6.2.2

## Durchführung: Usability-Test

In unserem Fall dient der Usability-Test dazu, herauszufinden, ob die Benutzer die Szenarien und die Fragestellungen der Umfrage verstehen, oder ob wir die Texte anpassen müssen. Zum Beispiel interessiert uns, ob die Wortwahl der Fragen sinnvoll ist und ihre Abfolge logisch ist. Des Weiteren wollen wir testen, ob die vier HTML-Prototypen einwandfrei funktionieren. Diese Ergebnisse helfen uns, die Prototypen vor der quantitativen Umfrage zu verbessern. Wir erhoffen uns auch, durch die Reaktionen der Benutzer herauszufinden, welchen der vier Prototypen sie bevorzugen und warum. Dazu lassen wir die Benutzer die verschiedenen Prototypen bewerten.

Im Verlauf des Projekts haben sich zudem noch spezifische Fragestellungen und zu überprüfende Punkte angesammelt. Diese stammen aus Ideen des Projektteams sowie den Anregungen der Auftraggeberin.

- Ist das beschriebene Szenario für die Testpersonen verständlich?
- Funktionieren die Prototypen einwandfrei?
- Welche Fortschrittsanzeige gefällt den Testpersonen am besten und was sind die Gründe dafür? Wir lassen die Benutzer die Fortschrittsanzeige nach bestimmten Kriterien ordnen.



Abbildung 6.3: Fortschrittsbalken mit Prozenten

- Erkennen die Testpersonen den Einstieg in die Shop-Seite via textlich beschriebenem Szenario?
- Ist die Positionierung der Fortschrittsanzeige ausreichend, so dass der Benutzer Aktualisierungen bemerkt und versteht? Die Positionierung der Fortschrittsanzeige ist bei allen 3 Prototypen mit Anzeige gleich. Wir wollen durch den Usability-Test herausfinden, ob die Benutzer die Anzeige bemerken.
- Sind die Fragen in einer logischen Abfolge?
- Sind die Fragen verständlich?

### Teilnehmer

Insgesamt nehmen vier Personen am Usability-Test teil (Abbildung 6.4). Alle sind UBS-Mitarbeiter, da sie ohne zusätzliche Bewilligungen anfordern zu müssen Zutritt zum UBS Usability-Lab bekommen können. Es nehmen drei Frauen und ein Mann am Test teil. Die Teilnehmer sind zwischen 24-58 Jahre alt und sind, wie im Kapitel 3.2.1 *Stakeholder Interview* bereits erwähnt, nicht im Besitz eines GAs. Sie stammen alle aus der Umgebung von Zürich und jeder ist ein potentieller Benutzer der SBB-Desktop-Applikation. Zwei der Teilnehmer arbeiten beruflich im UX-Bereich und sind mit dem Ablauf eines Usability-Tests vertraut. Durch ihre User Experience-Erfahrung können sie uns UX-relevantes Feedback zu den Prototypen liefern. Die beiden anderen Benutzer arbeiten im Bereich Software Localization, deren Arbeit darin besteht, GUIs, Formulare und Spezifikationen auf Inkonsistenzen, missverständliche Passagen und Verständnisprobleme zu prüfen. Sie erledigen teilweise auch Übersetzungen. Ihr sprachliches Wissen hilft uns enorm, das Wording des Fragebogens zu verbessern.

### Sprache

Der Usability-Test wird in den zwei Sprachen Englisch und Deutsch durchgeführt, da eine der Benutzerinnen aus England stammt. Wir übersetzen den Fragebogen ins Englische. Da die Benutzerin auch etwas Deutsch versteht, übersetzen wir zwar das Szenario, aber nicht die Eingabefelder des Prototypen. Benutzer fühlen sich wohler, wenn sie in ihrer Muttersprache Auskunft geben können, deswegen führen wir den Test in ihrer Muttersprache Englisch durch (Nielsen, 1996).

### Durchführung

Die Rollen in unserem Test sind Testleiter und Beobachter. Jedes Gruppenmitglied konnte während des Usability-Tests mindestens einmal jede Rolle übernehmen (siehe Kapitel 4.4).

Der Testleiter folgt einem Skript, damit alle vier Tests den gleichen Ablauf haben, übernimmt die Vorstellung aller anwesenden Personen, erklärt den Ablauf des Usability-Tests, stellt die Pre-Interview-Fragen (siehe Anhang 6), erklärt das Szenario und übernimmt die Nachbefragung.

Die Pre-Interview-Fragen befassen sich mit allgemeinen Daten wie Alter, Geschlecht, Online-Affinität, Reiseverhalten (wie oft wird Zug gefahren), Abonnement und häufigste Kaufmethode (siehe Anhang 7).

Diese Fragen helfen uns, die Benutzer und deren Bedürfnisse in Bezug auf die SBB besser kennenzulernen. Der Testleiter bittet den Teilnehmer, laut zu denken, damit wir so viel Feedback wie möglich erhalten. Er/sie weist ihn auch darauf hin, dass wir die Applikation testen und nicht sie selber und dass sie in diesem Usability-Test als Experten fungieren.

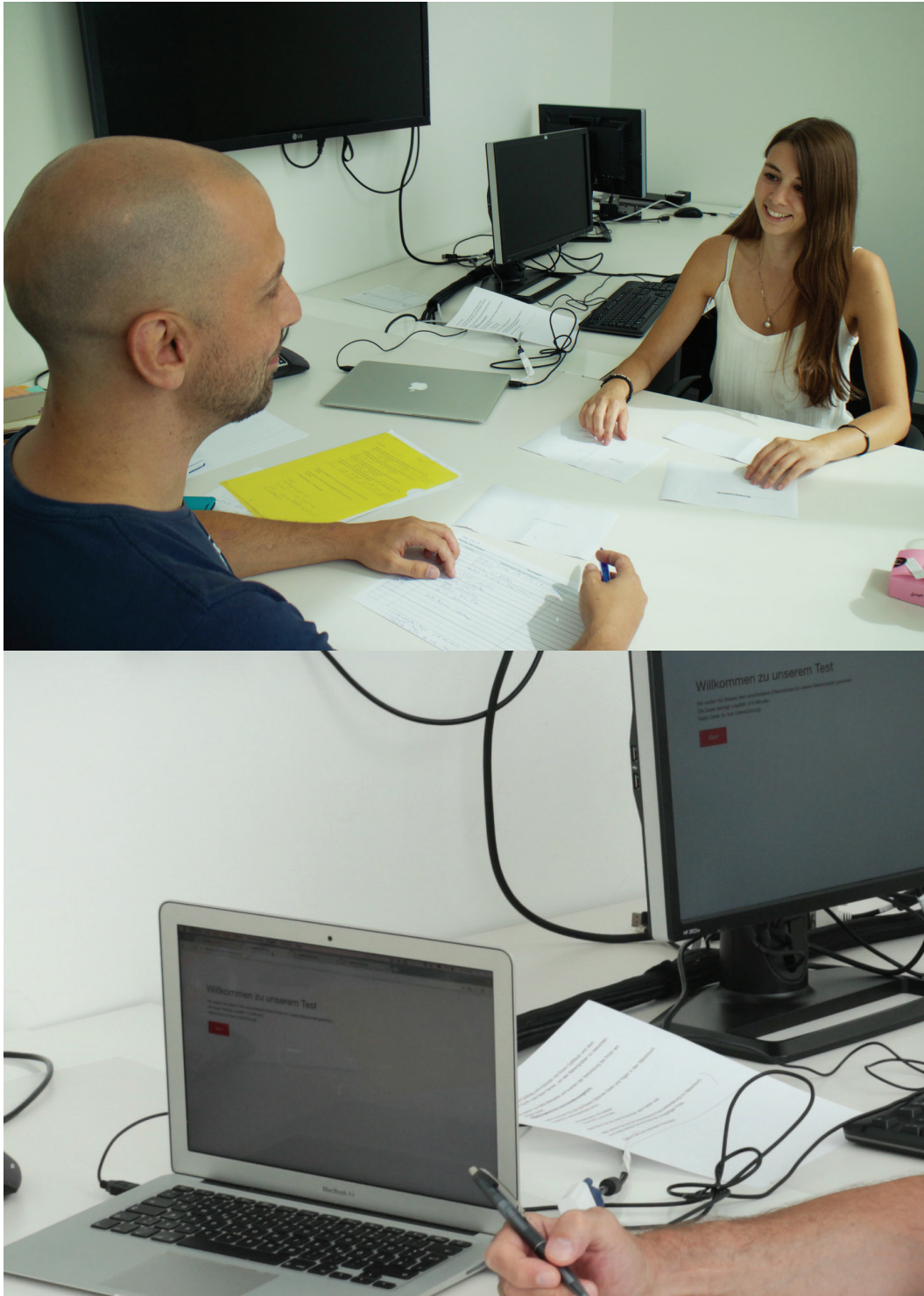


Abbildung 6.4: Fotos aus den Usability-Tests

Jede Testperson klickt sich durch jeden der vier Prototypen und füllt nach jedem Prototypen/Durchgang einen ausgedruckten Fragebogen aus. Dies steht im Unterschied zur Online-Umfrage, in der jeder Teilnehmer einen Prototyp zugewiesen bekommt. Jeder Person im Usability-Test werden die Prototypen in einer anderen Reihenfolge präsentiert. Zum Schluss stellen wir den Usability-Teilnehmern vier Fragen und bitten sie, die Fortschrittsanzeigen nach den unten erwähnten Kriterien zu ordnen.

- 1 Welchen Ladevorgang haben Sie als am langsamsten empfunden? Legen Sie eine Reihenfolge fest und begründen Sie Ihren Entscheid.
- 2 Welchen Ladevorgang haben Sie als am schnellsten empfunden? Legen Sie eine Reihenfolge fest und begründen Sie Ihren Entscheid.

3 Welcher Ladevorgang hat Ihnen am besten gefallen? Legen Sie eine Reihenfolge fest und begründen Sie Ihren Entscheid.

4 Wenn Sie heute eine Entscheidung fällen müssten, welchen Ladevorgang die SBB verwenden sollen, welchen würden Sie auswählen?

Diese Art der Befragung kann nur in einem Usability-Test durchgeführt werden, da ein direkter Austausch mit den Benutzern stattfindet. Wir erhoffen uns, dadurch Antworten zur allgemeinen Zufriedenheit über die Ladedauer und die visuelle Attraktivität der Fortschrittsanzeige zu erhalten, die im Kapitel 5.1 definiert ist.

Zum Schluss fragen wir nach weiteren guten Beispielen von Fortschrittsanzeigen und Gründen dafür, dass sie als gute Lösungen empfunden werden. Wir notieren die Meinungen und Feststellungen zum Thema Ladedauer (Abbildung 6.5).

Usability Räumlichkeiten

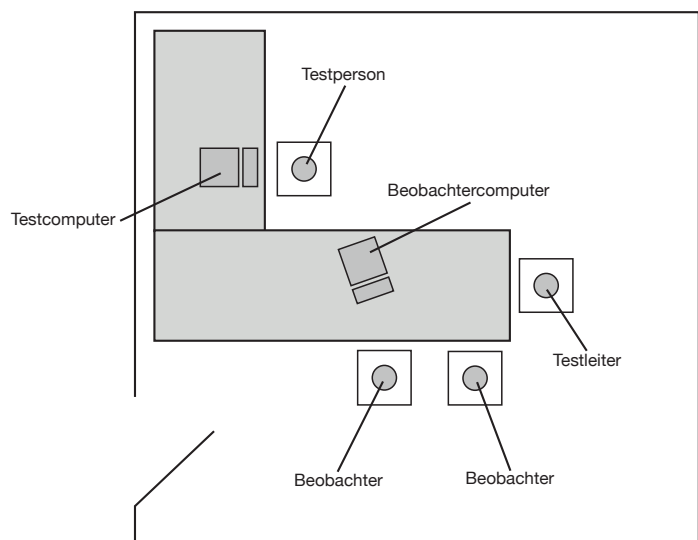


Abbildung 6.5: Usability Räumlichkeiten

### 6.2.3

## Nachbearbeitung: Erkenntnisse aus den Usability-Tests und dem Hallway Testing

Um die Ergebnisse des Usability-Tests zu evaluieren, bearbeitet jedes Gruppenmitglied die Ergebnisse von ein bis zwei Testpersonen, liest deren Antworten auf die Fragen, notiert Zusatzkommentare und schlussendlich, wie sie die Fortschrittsanzeigen bewerteten. Wir tragen alle Ergebnisse zusammen und schauen, welche Findings sich decken und tragen sie alle in einer Excel-Tabelle ein. Für die Nachbearbeitung der Erkenntnisse aus den Usability-Tests benutzen wir die Einstufungsskala von Nielsen (1995). Mit Hilfe dieser Skala priorisieren wir die gefundenen Probleme. Niensens Skala hat 5 Stufen und geht von 0 (kein Usability-Problem) bis 4 (Usability-Katastrophe). Diese Skala ist sehr verbreitet und wir sind mit ihrer Benutzung vertraut. Ein Beispiel unserer Priorisierung kann man im Anhang 8 finden.

Die wichtigsten Erkenntnisse und die nachfolgenden Korrekturen aus den Usability-Tests lauten wie folgt:

Abgesehen von wenigen textlichen Änderungen sind die Fragen des Fragebogens verständlich formuliert und ihre Abfolge ist klar definiert.

- Wir bestätigen, dass die Ladeanzeige auf dem Warenkorb für den Benutzer sichtbar ist.
- Ein paar Bugs in den Prototypen wurden gefunden und entsprechend korrigiert. Wir bekommen gutes Feedback für die Prototypen.

- Der Einstieg zur Reiseoptionenseite ist für die Benutzer nicht klar genug definiert, da das Szenario textlich beschrieben ist. Sie haben das statische Bild des Fahrplans nicht verstanden und wussten nicht, wie sie zur Reiseoptionenseite navigieren sollen. Ein bildliches Storyboard würde das Szenario und den Ablauf für den Benutzer klarer machen und ihn direkt zur Reiseoptionenseite führen (Abbildung 6.6).
- Das Interaction Design können wir dadurch verbessern, dass der Benutzer jetzt nach der Eingabe via Enter-Taste zum nächsten Feld springen kann.
- Ein Bug in den Prototypen wird erkannt und korrigiert: Die Eingabe der Input-Felder wurde, während der Benutzer einen Reisenden löscht, nicht geblockt.

Die Benutzer fanden die Prozentangaben im Fortschrittsbalken nicht notwendig und ihr Feedback ist, dass es besser sei, ein Fortschrittsbalken ohne Prozentangaben als Fortschrittsanzeige abzubilden.

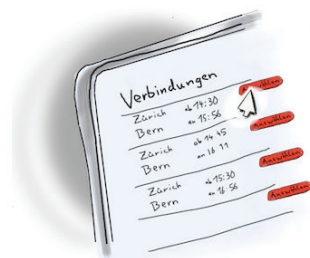
Wir erwarteten, dass alle Testbenutzer den gleichen Prototypen als ihren Favoriten auswählen würden, waren jedoch überrascht, dass von den vier Personen, die wir befragt haben, je zwei den gleichen Prototypen auswählten: Prototyp 3 (Spinner mit Text) und 4 (fahrender Zug), die sich sehr voneinander unterscheiden. Gründe dafür sind, dass ihnen die Art und Weise des Ladevorgangs bei Prototyp 3 bekannt war (Spinning Wheel) und die zusätzlichen Informationen hilfreich sind. Allerdings sollten sie langsam eingeblendet werden, damit der Benutzer eine Chance hat, sie zu lesen.

## Stell dir vor...

Du hast heute frei und möchtest am Nachmittag mit deinem Neffen und dem Grossvater von **Zürich nach Bern** fahren, um die Stadt zu besuchen.



Du gehst zur SBB-Website und gibst deine Suchanfrage für die Verbindung ein.



Nun klickst du auf eine der angezeigten Verbindungen um weiter zu kommen.

Weiter ...

Abbildung 6.6: Storyboard

Eine weitere Erkenntnis ist die, dass der Prototyp 4 (fahrender Zug) bei den anderen beiden Benutzern zwar gut angekommen ist, sie aber den vorbeifahrenden Zug nach einer Weile als etwas mühsam empfanden und deshalb Zug-Variationen vorschlugen.

### Hallway Testing

Hallway Testing, oder auch Guerrilla Testing genannt, ist eine Methode, mit der die Rekrutierung von Testbenutzern am einfachsten ist. Es können Benutzer in ihrer gewohnten Umgebung befragt werden, zum Beispiel Arbeitskollegen im Büro, Freunde, Leute auf der Strasse und im Cafe. Der Hallway Test sollte kurz gehalten sein und deshalb nicht mehr als 15 Minuten dauern. Zwar können keine kompletten Szenarios getestet werden wie bei einem Usability-Test, aber erste Usability-Stolpersteine können erkannt und korrigiert werden. Der Nachteil dieser Methode ist, dass das Benutzerprofil nicht gesamtheitlich auf die Benutzergruppe abgestimmt ist. Das heisst, ihr Feedback ist nicht 100% repräsentativ (Hübscher, 2015).

Wir führen Hallway Tests mit zwei Personen durch. Die beiden Tests konzentrieren sich nur auf den Ablauf des Prozesses, das visuelle Erscheinungsbild der Prototypen (Abbildung 6.7), die Verständlichkeit des Fragebogens und nicht auf die Umgebung, da sie in unserem Fall zu Hause mit unseren Partnern stattfinden.

Das Ergebnis ist, dass die Prototypen einwandfrei funktionieren und der Benutzer ohne Verzögerung am Ende auf den Fragebogen weitergeleitet wird. Das Szenario wird durch das Storyboard in Bildern gut verstanden, besser als das textliche Szenario im Usability-Test.

## 6.3

### Hauptversuch: Online-Umfrage

Als Hauptteil unserer Studie führen wir eine quantitative Analyse durch, wie im Kapitel 4 erklärt. Dafür gestalten wir einen Fragebogen, der zusammen mit den Prototypen in einer Online-Befragung benutzt wird.

#### 6.3.1

### Fragebogen

Ein Fragebogen eignet sich sehr gut für die Datenerhebung in einer quantitativen Studie, da die Fragen standardisiert und für alle Befragten gleich sind wie auch die Antwortmöglichkeiten, die statistisch ausgewertet und verglichen werden können (Courage/Baxter, 2005: 313f.).

Unser Ziel ist, mit dem Fragebogen die Zufriedenheit und Aussagekraft der Fortschrittsanzeigen zu messen. Der Fragebogen wird an den Prototypen angehängt, um die Datenerhebung via Online-Umfrage durchführen zu können.

Es gibt auf dem Markt verschiedene standardisierte Fragebogen, die speziell für die Messung der User Experience oder der Usability einer Applikation entwickelt wurden. Da der Kern unserer Arbeit in diesem Bereich liegt, evaluieren wir nachfolgende Optionen auf ihre Tauglichkeit für unser Projekt. In Frage kommen die uns bekannten QUIS- und SUS-Fragebogen, wie auch der S-UEQ von SAP.

Der QUIS-Fragebogen (kurz für *Questionnaire for User Interaction Satisfaction*) wurde an der Universität von Maryland in den USA entwickelt.

Der QUIS-Fragebogen misst anhand einer Likert-Skala neun verschiedene Faktoren, die die User Experience beeinflussen. Diese Faktoren decken ein breites Spektrum eines Produkts ab, von Terminologie, Hilfe und Handbüchern bis zur Installation und Multimedia-Fähigkeiten (Sauro/Lewis, 2016: 189f.; Albert/Tullis, 2010: 139f.).

SUS (kurz für *System Usability Scale*, Brooke, 1996) ist ein Fragebogen mit 10 Fragen, der im Jahr 1986 von John Brooke entwickelt wurde. SUS ist sehr verbreitet, da er eine einfache, günstige und standardisierte Lösung bietet, um die Usability eines Produkts zu messen. Der SUS-Fragebogen ist mit den Jahren zum Industriestandard geworden und bietet dank der grossen Verbreitung eine optimale Vergleichsmöglichkeit (Sauro, 2011).

Der S-UEQ (kurz für *Short User Experience Questionnaire*, Laugwitz et al., 2008) wurde von Laugwitz, Schrepp und Held im Jahr 2005 entwickelt. Der Fragebogen ist eine einfache Liste von 26 Adjektiven zusammen mit ihren Gegensätzen. Die Adjektive werden in 6 Gruppen für die Auswertung zusammengefasst. Dieser Fragebogen eignet sich gut für das Testen der User Experience von Produkten, und seine Effektivität ist validiert (UEQ Handbook, Schrepp, 2015).

Wir entscheiden uns aus folgenden Gründen gegen die Benutzung eines dieser Standard-Fragebogens:

- Unser Ziel für die Arbeit ist sehr konkret auf die Fortschrittsanzeigen und die Benutzerzufriedenheit fokussiert. Die Standardfragebogen sind zu allgemein gehalten und dienen unserem Ziel zu wenig.
- Die Benutzung des QUIS-Fragebogens ist ausserdem kostenpflichtig.
- Der Einsatz von S-UEQ setzt spezifische Erfahrung voraus, über die wir im Team nicht verfügen.

Deshalb beginnen wir mit dem Design unseres eigenen Fragebogens.

Um unsere Hypothesen zu validieren, benötigen wir Fragen für die Messung der Zufriedenheit, Angemessenheit und Aussagekraft der Fortschrittsanzeigen in den Prototypen.

Wir beginnen Mitte August mit der Entwicklung des Fragebogens in mehreren Workshops und Brainstormings und verfeinern die Ergebnisse mit Hilfe von Reviews durch Experten und unabhängigen Personen (Abbildung 6.8). Diese Arbeit ist anspruchsvoll und sehr zeitintensiv und stellt einen äusserst kritischen Teil der Arbeit dar.

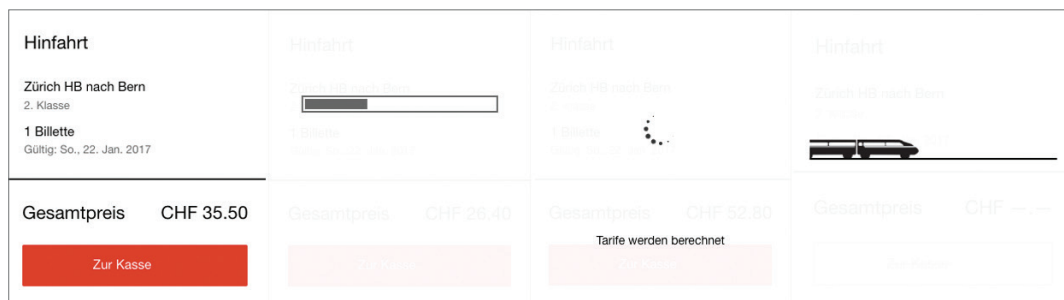


Abbildung 6.7: Finale Version der Fortschrittsanzeigen

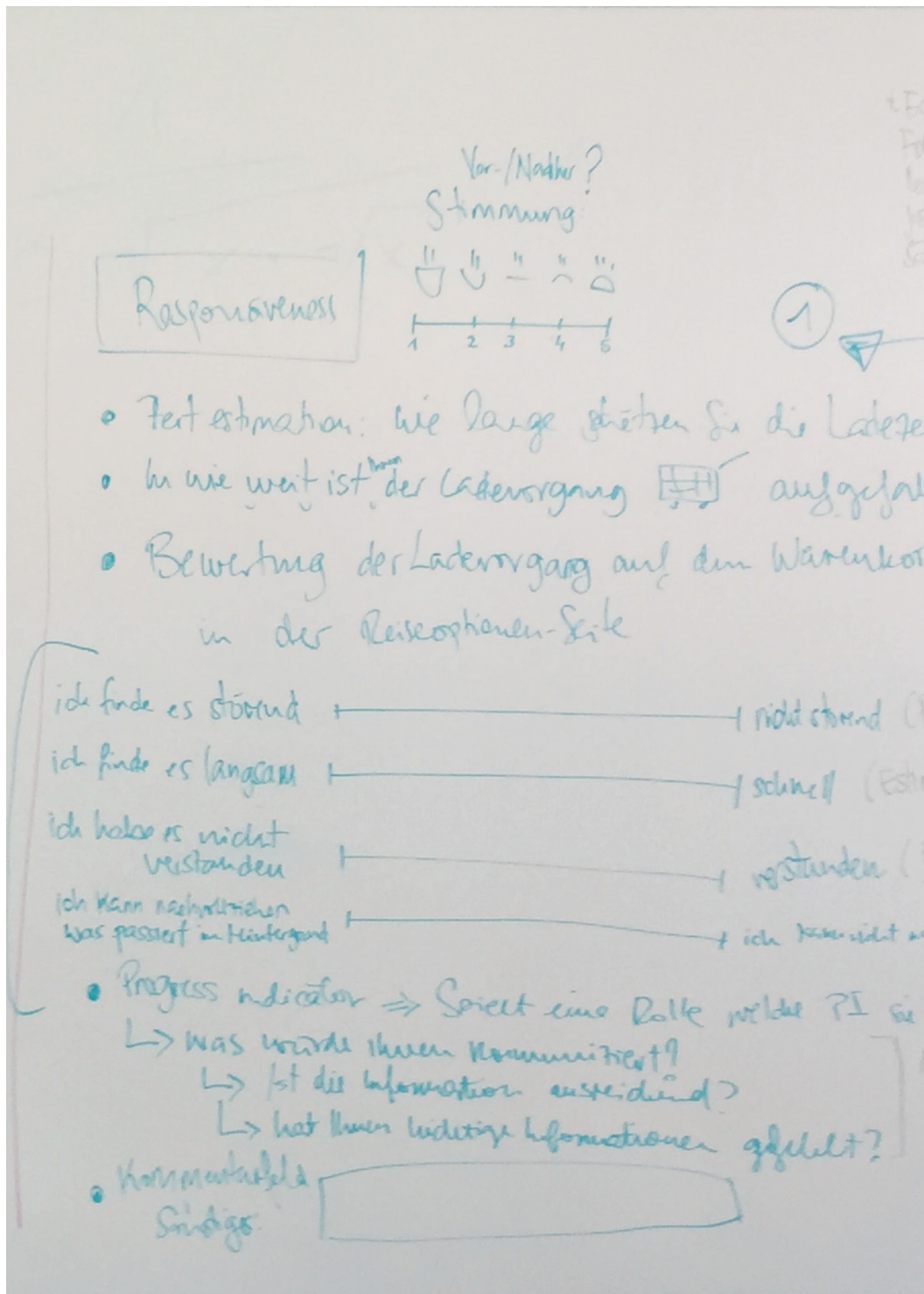


Abbildung 6.8: Workshop-Resultate für die Entwicklung des Fragebogens

Ein Beispiel ist der Begriff für die Fortschrittsanzeige. Wir starten mit dem Fachbegriff *Progress Indicator*, der für die IT-affinen Personen zwar ziemlich klar ist, für andere Personen jedoch geklärt werden müsste. Wir arbeiten dann mit dem deutschen Begriff *Ladevorgangsanzeige*. In verschiedenen Reviews mit unterschiedlichen Personen erhalten wir das Feedback, dass das Wort zu lang ist. Wir wählen dann *Anzeige des Ladevorgangs*, was wir dann auch im Fragebogen verwenden.

Auch das Messen der User Experience mit Fragen stellt uns vor eine Herausforderung. Wir testen verschiedene Fragetypen auf deren Wirkung, lassen uns dann vom S-UEQ Fragebogen inspirieren und arbeiten mit gegensätzlichen Adjektiven (Abbildung 6.9). Für die anderen Fragen arbeiten wir mit Sätzen und Bewertungstabellen.

Wir arbeiten mit vier unterschiedlichen Prototypen und brauchen Daten für jeden Prototyp, deswegen entscheiden wir uns, vier gleiche Fragebogen zu kreieren, von denen je einer an den entsprechenden Prototyp angehängt wird. Deswegen führen wir eine erste Kontrollfrage ein, auf die der Befragte antworten muss, was für Fortschrittsanzeige er gesehen hat (Abbildung 6.10).

Eine Kopie des Fragebogens befindet sich im Anhang 4.

### 6.3.2

## Durchführung der Online-Umfrage

Die technische Herausforderung ist hier, den Link zu einem Fragebogen-Anbieter derart zu gestalten, dass die Informationen, welcher Prototyp bearbeitet worden war, weitergegeben werden kann. Das stellte sich aber als sehr aufwändig heraus, so dass wir eine Alternative finden müssen. Wir entscheiden uns für eine einfache Lösung: wir legen 4 Fragebogen mit demselben Inhalt an, von denen je einer an einen Prototypen angehängt wird. So können wir durch vier verschiedene Verlinkungen nachvollziehen, welcher Prototyp bearbeitet worden ist.

Für die Durchführung der Online-Umfrage evaluieren wir verschiedene Umfrage-Services, die auf dem Markt existieren. In die engere Wahl kommen drei Tools:

- Google Forms
- Survey Monkey
- Onlineumfrage.ch

Survey Monkey ist schnell aus dem Rennen, da der Datensatz sich mit dem Gratis-Account nicht exportieren lässt und das Tool auf 10 Fragen und 100 Antworten limitiert ist.

Hast du eine Anzeige dieses Ladevorgangs gesehen?  
Falls ja, welche? (siehe Abbildung unten) \*

A  
 B  
 C  
 Nein, ich habe keine Abbildungen gesehen

<p>Hinfahrt <b>A</b></p> <p>Zug nr. 100   10:00</p> <p>100%</p> <p>Gesamtpreis CHF 10.00</p> <p>100%</p>	<p>Hinfahrt <b>B</b></p> <p>Zug nr. 100   10:00</p> <p>100%</p> <p>Tarife werden berechnet</p> <p>Gesamtpreis CHF 10.00</p> <p>100%</p>	<p>Hinfahrt <b>C</b></p> <p>Zug nr. 100   10:00</p> <p>100%</p> <p>Gesamtpreis CHF 10.00</p> <p>100%</p>
----------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------

Abbildung 6.10: Kontrollfrage im Fragebogen

Hypothesen	Fragen und Adjektive aus dem Fragebogen
<b>Hypothese 1</b> Wahrnehmung der Ladedauer	<ul style="list-style-type: none"> <li>• Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?</li> <li>• Ich empfand die Dauer als angemessen.</li> <li>• Wie hast du die Anzeige des Ladevorgangs empfunden?</li> <li>• Schnell/Langsam</li> </ul>
<b>Hypothese 2</b> Visuelle Ästhetik	<ul style="list-style-type: none"> <li>• Wie hast du die Anzeige des Ladevorgangs empfunden?</li> <li>• Attraktiv/Unattraktiv</li> <li>• Originell/Konventionell</li> <li>• Sympathisch/Unsympathisch</li> <li>• Spannend/Langweilig</li> <li>• Ich empfand die Darstellung als optisch attraktiv</li> </ul>
<b>Hypothese 3</b> Benutzerzufriedenheit mit der Ladedauer	<ul style="list-style-type: none"> <li>• Wie hast du die Anzeige des Ladevorgangs empfunden?</li> <li>• Nützlich/Nutzlos</li> <li>• Übersichtlich/Verwirrend</li> <li>• Angenehm/Lästig</li> <li>• Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen</li> </ul>
<b>Hypothese 4</b> Aussagekraft der Ladeanzeige	<ul style="list-style-type: none"> <li>• Wie hast du die Anzeige des Ladevorgangs empfunden?</li> <li>• Voraussagbar/Unberechenbar</li> <li>• Verständlich/Unverständlich</li> </ul>

Abbildung 6.9: Mapping Fragen und Adjektiven zu den Hypothesen

Onlineumfrage.ch gewinnt gegenüber Google Forms, weil die Firma auf Umfragen spezialisiert ist und einen Studenten-Account mit den wichtigsten Features für unsere Arbeit anbietet: Export der Daten, alle Fragetypen, unlimitierte Anzahl von Fragebögen.

### **Teilnehmer**

Für die Desktop-Applikation haben die SBB ein Panel von 500 Personen zusammengestellt (siehe Kapitel 3.2.1). Die Benutzung dieses Panels wäre eine gute Möglichkeit gewesen, da die Teilnehmer Desktop-User sind und das gleiche Profil haben wie wir es für unseren Usability-Test benutzen. Wie bereits erwähnt konnten wir nicht auf das Panel der SBB zugreifen.

Allerdings müssen die Teilnehmer in unserem Experiment nicht unbedingt über ein besonderes Profil verfügen. Zum einen ist so gut wie jeder Einwohner der Schweiz ein potentieller Benutzer des SBB-Ticketshops. Zum anderen fragen wir in unserem Experiment nur grundlegende menschliche Reaktionen und Emotionen ab, die relativ unabhängig vom jeweiligen Individuum sind. Zum Beispiel sollten die Unterschiede in der Wahrnehmung des Ladeverhaltens, das innerhalb weniger Sekunden geschieht, nicht allzu sehr von individuellen Eigenschaften abhängen. Zudem wird durch eine möglichst hohe Anzahl an Teilnehmern dafür gesorgt, dass sich eventuell vorhandene Unterschiede ausgleichen.

Um eine statistisch relevante Anzahl Teilnehmer zu erhalten, haben wir uns auf unser privates Netzwerk verlassen. Jedes Gruppenmitglied hat circa 100-150 E-Mail-Adressen gesammelt. Wir bekamen Unterstützung von Freunden, Arbeitskollegen und Familienmitgliedern, die entweder

die E-Mails weiterleiteten oder uns Adressen aus ihrem privaten Netzwerk zur Verfügung stellten.

### **Vorankündigung**

Um die Rücklaufquote zu erhöhen, wird ein Vorankündigungsemail verschickt. Eine Vorankündigung ist eine erste Kontaktaufnahme mit dem Teilnehmer. Basierend auf dem Artikel von Klein/Rolf Porst (2000: 17), gilt generell, dass jeder zusätzliche Kontakt eine Rücklaufquote verbessert.

Vorankündigungen sollen jedoch nur dann verwendet werden, wenn die Umfrage ein einziges Mal verschickt wird. Hätten wir eingeplant, den potentiellen Teilnehmern, die die Umfrage versäumt haben, ein Nachfass-E-Mail zu schreiben, hätte sich ein erster Kontakt erübrigt. In unserem Fall ist dieses E-Mail angebracht, da die meisten Kontakte aus unserem privaten Netzwerk stammen (Courage/Baxter, 2005: 318).

### **Randomisierung**

Falls bei einem Quasi-Experiment mittels Randomisierung kein Einfluss auf die Teilnehmer-Gruppe genommen wird und die gesammelten Teilnehmer zufällig diesem Experiment zugeteilt werden, die Liste in keiner Weise sortiert wurde und somit alle Teilnehmer die gleichen Merkmale besitzen, spielt es keine Rolle, in welcher Reihenfolge die Prototypen verschickt werden. Dies bedeutet für unsere Arbeit, dass Prototyp 1 zum Beispiel an die ersten 100 Teilnehmer hätte verschickt werden können. Danach hätten Prototyp 2, 3 und 4 auf die gleiche Art verteilt werden können. In unserem Fall gehen wir davon aus, dass alle Teilnehmer die gleichen Merkmale besitzen und wir haben die Teilnehmerliste zufällig zusammengestellt (Minitab 17, 2016).

### Durchführung

Die potenziellen Teilnehmer sollen rein zufällig einem der vier Prototypen zugeordnet werden. Ein einzelner Teilnehmer darf nur einen einzigen Prototyp bearbeiten, unbeeinflusst von allen anderen Prototypen. Damit können wir alle Prototypen neutral miteinander vergleichen (Randomisierung). Das bedeutet, die Prototypen werden in der Reihenfolge 1 - 4 verschickt (Abbildung 6.11).

Wir haben die Fragebogen folgendermassen im Monat Oktober verschickt:

Wir starten mit 4 Teilnehmern pro Tag und steigern die Teilnehmerzahl am 5. Tag auf 84. Am Anfang füllt jeder 5. Teilnehmer den Fragebogen aus. Dies steigert sich nach und nach. Einige Teilnehmer weilen auch in den Ferien und nehmen erst später an der Umfrage teil. Wir haben mit den Prototypen 2, 3 und 4 schnell

Fortschritte gemacht. Im Laufe des Monats bleibt die Teilnehmerrate des Prototypen 1 am niedrigsten und wir vermuten eine zu hohe Ausstiegsrate. Der Grund, so unsere Annahme, liegt darin, dass keine Fortschrittsanzeige angezeigt wird. Für 7 Sekunden passiert nichts, die Teilnehmer wissen nicht, was sie sehen sollen, vermuten einen Fehler und brechen die Umfrage frühzeitig ab. Um die gleiche Anzahl Teilnehmer zu erreichen, senden wir an die restlichen Adressen am Ende des Monats den Link zum Prototypen 1. Sobald die Prototypen 2, 3 und 4 eine Teilnehmerzahl von 70 erreicht haben, werden die Teilnehmer automatisch zum Prototypen 1 weitergeleitet.

Insgesamt haben wir circa 500 E-Mails verschickt. Da unsere Bekannten und Freunde uns tatkräftig halfen, wissen wir nicht genau wieviele zusätzliche E-Mails sie aus eigenem Ansporn verschickten.

Tage	1	2	3	4	5	6	7	8	9	10
Anzahl Teilnehmer	4	11	57	84	98	86	48	5	37	15

Abbildung 6.11: Anzahl versendeter E-Mails pro Tag



# 7. Ergebnisse

Sobald die Online-Umfrage beendet ist, starten wir in die Datenanalyse-Phase. Da wir im Team nicht über genug statistisches Know-How verfügen, suchen wir uns extern Hilfe in Form einer Statistik-Masterstudentin.

Sie berät uns bei der Auswahl der Testverfahren und führt später die Berechnungen durch, da wir kein geeignetes statistisches Programm besitzen.

Die Berechnungen wurden mit Statistica (Version 13) von statsoft erstellt. Die Power-Kalkulation und die Anpassungen für multiples Testen wurden mit der Programmierumgebung R 3.3.2 (freeware), mit dem Add-on *pwr* erstellt.

## 7.1

### Überarbeitung des Datensatzes

Sobald wir bei der Online-Umfrage eine akzeptable Anzahl Antworten erreicht haben, konsolidieren wir die Datensätze, die *onlineumfrage.ch* uns liefert, in einem einzigen Datensatz für die weitere Bearbeitung. Dafür fügen wir eine Spalte ein, in die die Nummer des zugehörigen Prototypen eingetragen wird. Zu diesem Zeitpunkt haben wir 282 Antworten.

Im konsolidierten Datensatz entfernen wir zuerst die Zeilen der Fragebogen, in denen die Befragten unsere Kontrollfrage falsch beantwortet haben. Das sind 8 Antworten (3 aus dem Prototyp 1, 1 aus dem Prototyp 2, 3 aus dem Prototyp 3 und 1 aus dem Prototyp 4).

Der nächste Schritt ist, den Datensatz nach weiteren Inkonsistenzen in den Antworten zu durchsuchen. Konkret betrachten wir die Frage *Während des Ladevorgangs hat sich der Warenkorb aktualisiert. Hast du diese Veränderung*

*bemerkt?* und die nachfolgende Frage *Wenn ja, hat sich der Warenkorb so verändert wie du es erwartet hast?*. Wir schauen uns die Antworten an, die nicht sinnvoll sind. Das sind Einträge von Teilnehmern, die die Änderung im Warenkorb nicht bemerkt haben, aber trotzdem antworten, dass die Änderung ihren/seinen Erwartungen entspricht oder nicht.

In diesem Fall haben wir 30 Einträge, die wegen Inkonsistenz verdächtig sind. Wir prüfen die Antworten auf früheren Fragen, besonders die Antworten auf die Fragen:

- Wie hast du die Anzeige des Ladevorgangs empfunden?
  - unverständlich (1) - verständlich (7)
  - verwirrend (1) - übersichtlich (7)
- Im Vergleich mit deinen sonstigen Erfahrungen im Internet, wie bewertest du den Ladevorgang auf der Testseite?
  - Ich konnte nachvollziehen, wieso diese Anzeige des Ladevorgangs erschien.

Nach Überprüfung dieser 30 Fälle eliminieren wir noch 2 weiteren Einträge.

Bei der Berechnung der Statistik schliesst unsere Statistikerin noch 3 Fälle von der Berechnung aus, da die Antworten auf die Frage *Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?* unrealistisch waren wie 654'754'563'653, 9'999 oder 1 Milliarde. Diese 3 Fälle werden trotzdem in der deskriptiven Statistik berücksichtigt, unter Ausschluss der entsprechenden Frage.

Nach der Bearbeitung arbeiten wir mit einem Datensatz mit 269 Fällen.

## 7.2

### Deskriptive Daten

Unsere Stichprobe von 269 Elementen kann demographisch wie folgt beschrieben werden:

63% sind Männer und 37% Frauen, was ein Verhältnis von 1 zu 1.72 bedeutet. Etwas mehr als die Hälfte der Befragten (53%) befindet sich in der Altersgruppe von 26 bis 40. In dieser Altersgruppe ist das Verhältnis Mann/Frau mit 1.42 etwas besser. Personen über 56 (7%) und Jugendliche unter 25 (4%) haben die geringste Repräsentation (Abbildung 7.1).

80% der Stichprobe sehen sich als technikaffin, sie beantworteten die Fragen nach der Technologie-Affinität mit *Trifft absolut zu* (siehe Anhang 1, S. 22).

Ungefähr 3 Viertel der Stichprobe (73%) fahren regelmässig Zug: von *täglich* (34%) bis zu mehrmals pro Woche (17%) oder mehrmals pro Monat (22%). Nur 27% fahren selten oder nie Zug (Abbildung 7.2).

Beim Kaufen der Tickets bevorzugen unsere Befragten die App auf dem Smartphone mit 38% Prozent der Antworten (Abbildung 7.3). Danach kommt der Billettautomat (18%) und SBBs Webshop (17%). Von den SBB wissen wir, dass die meisten Billette am Automaten verkauft werden, gefolgt von der App und dann dem Webshop (siehe Stakeholder Interview, Kapitel 3.2.1). Wir haben daher in unserer Stichprobe eine davon abweichende Verteilung erhalten.

Für die Überprüfung der Hypothesen sind die Antworten gleichmässig auf die vier Prototypen verteilt, das heisst, wir haben 25% der Antworten für jeden Prototyp.

Der Durchschnitt der wahrgenommenen Ladezeit nach Prototyp sieht wie folgt aus (in dieser Berechnung sind die Ausreisser-Werte ausgeschlossen, siehe 7.1 - Abbildung 7.4):

Unsere Befragten empfanden den Prototypen 3 mit der korrekten Anzeige gemäss Literatur (Seow, 2008; Mishunov, 2015) für die 7-Sekunden Wartezeit als den langsamsten, und den Prototyp 2 als den schnellsten. Der Unterschied ist jedoch statistisch nicht signifikant (siehe Kapitel 7.3).

Die Fragen mit den Antworten *Ich empfand die Dauer als angemessen* und *Die Anzeige half, die Wartezeit kürzer zu machen* zeigen eine Tendenz zu *nicht treffend*. 68% der Befragten fanden, dass die Wartezeit nicht angemessen war, und die Anzeige half 57% der Befragten nicht, die Wartezeit kürzer zu machen.

Die Antworten für *Ich konnte nachvollziehen, wieso die Anzeige der Ladevorgangs erschien* und *Ich empfand die Darstellung als optisch attraktiv* sind anders verteilt. 49% konnten nachvollziehen, warum sie eine Anzeige gesehen haben, und 35% fanden die Anzeige schön.

In der Frage mit den gegensätzlichen Adjektiven finden wir in der Häufigkeiten- und Zentralwert-Berechnung bei 8 Adjektiven eine klare Tendenz zur mittleren Antwort (Abbildung 7.5).

Die Verteilung des Zentralwerts für die Adjektive pro Prototyp sieht man in der Abbildung 7.6.

Prototyp 2 mit dem Fortschrittsbalken erhält höhere Werte bei *nützlich*, *voraussagbar* und *verständlich*, wahrscheinlich wegen der grösseren

Bekanntheit der gewählten Fortschrittsanzeige. Prototyp 4 schnitt in fast allen Adjektiven besser als die anderen Prototypen ab, vor allem bei attraktiv, originell, angenehm und spannend. Die Frage mit den Adjektiven bezieht sich auf die Ladeanzeige. Prototyp 1 (ohne Fortschrittsanzeige) schnitt in fast allen Kriterien schlechter ab als die anderen Prototypen.

Weitere deskriptive Daten aus unserer Studie befinden sich im Anhang 1.

Prototyp	Mean
Prototyp 1	6.00
Prototyp 2	5.41
Prototyp 3	6.39
Prototyp 4	5.43

Abbildung 7.4: Durchschnitt der wahrgenommenen Ladezeit pro Prototyp

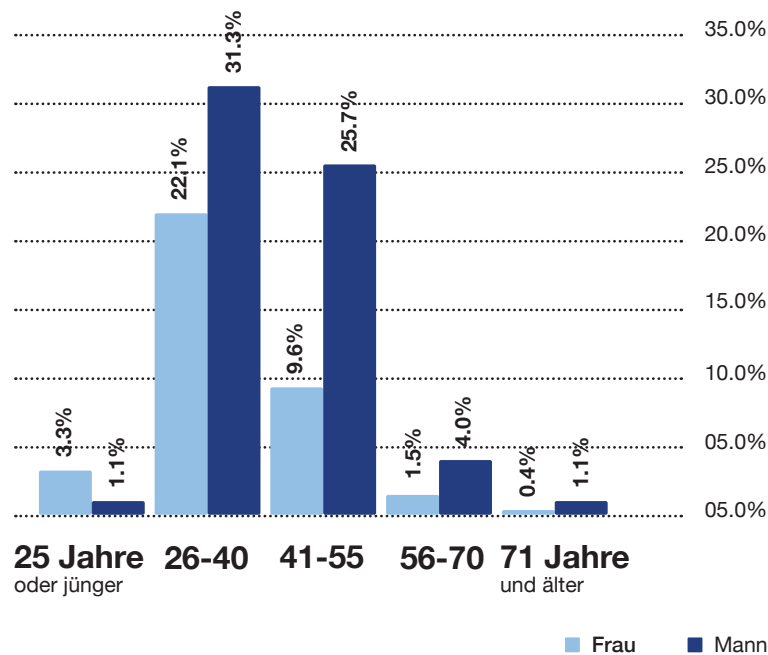


Abbildung 7.1: Geschlecht und Alter

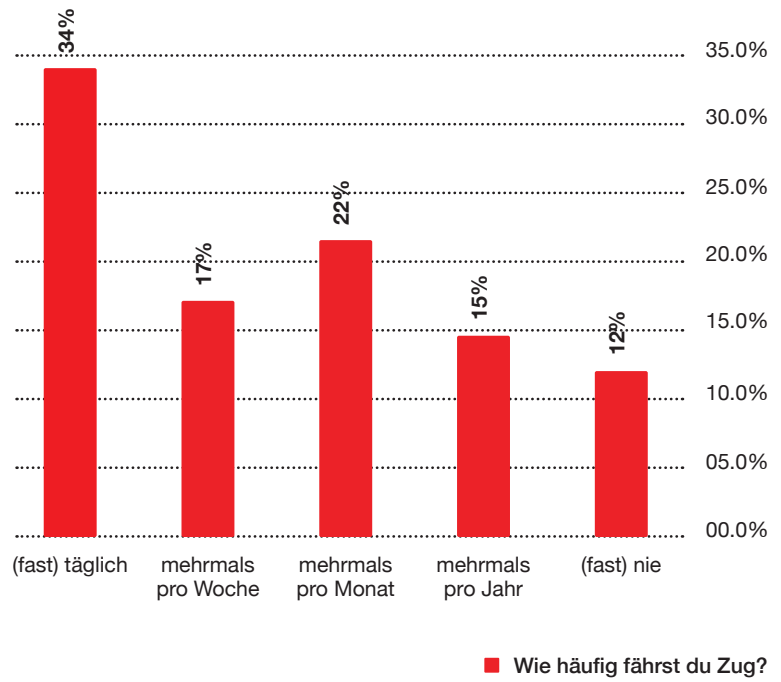


Abbildung 7.2: Zugfahrverhalten

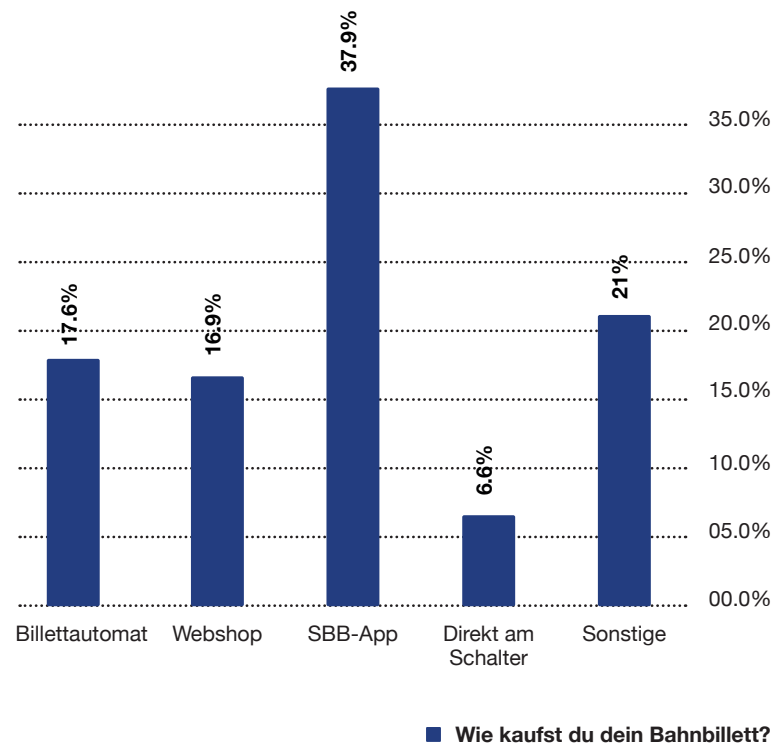


Abbildung 7.3: Billettkauf-Optionen

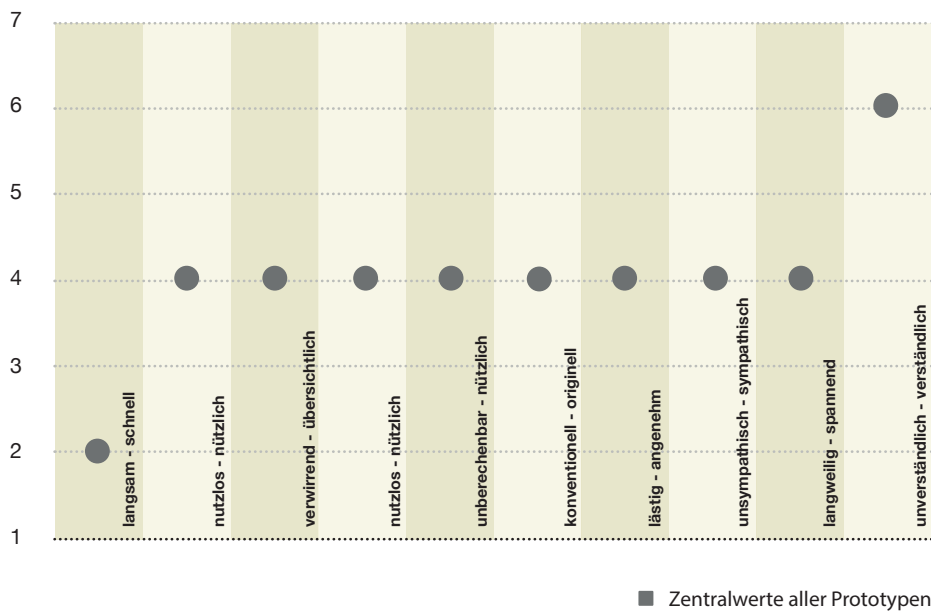


Abbildung 7.5: Zentralwert-Berechnung aller Prototypen

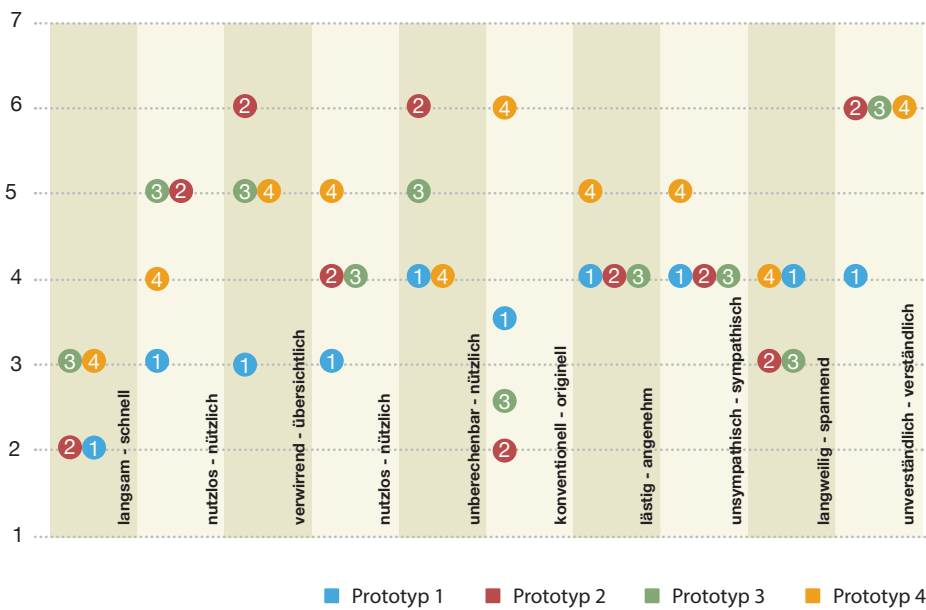


Abbildung 7.6: Zentralwert-Berechnung pro Prototyp

### 7.3

## Validierung der Hypothesen

In diesem Kapitel beschreiben wir die Ergebnisse der Hypothesen, die in Kapitel 8 interpretiert werden. Die Berechnungen befinden sich im Anhang 2.

### 7.3.1

## Der Einfluss der Fortschrittsanzeige auf die empfundene Ladedauer hat sich nicht bestätigt

Wie in Kapitel 5.3 beschrieben, stellt die 1. Hypothese die Behauptung auf, dass **eine Ladeanzeige hilft, dass der Benutzer die Ladedauer als kürzer einschätzt, bzw. wahrnimmt.**

Um diese Hypothese zu testen, wurde zunächst überprüft, ob es signifikante Unterschiede gibt zwischen den 4 Prototypen bezüglich der Frage: *Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?*

Die Auswertung wurde mit einem Kruskal-Wallis-Rangsummentest vorgenommen. Dieser ergibt einen p-Wert von **0.6994** (siehe Anhang 2, S. 2). Dies ist weit entfernt von statistischer Relevanz (definiert mit der üblichen Signifikanzgrenze von **0.05**). Hieraus können wir schliessen, dass es für die Dauer die man schätzt keinen Unterschied macht, welchen Prototypen, bzw. welche Anzeige, man gezeigt bekommt. Die Auswertung von Prototyp 1 (ohne Fortschrittsanzeige) im Vergleich zu allen anderen Prototypen (mit Fortschrittsanzeige) verdeutlicht diese Schlussfolgerung (Mann-Whitney-U-Test, p-Wert **0.6344**, siehe Anhang 2, S. 3).

Um alle vier Prototypen miteinander zu vergleichen und den Vergleich von den Prototypen mit einer Fortschrittsanzeige und dem einen ohne Anzeige für die Frage *Ich empfind die Dauer als angemessen* zu überprüfen, wenden wir den Pearson Chi-Quadrat-Test an.

Das Resultat ergibt einen p-Wert **0.17884** (siehe Anhang 2, S. 6) (der Vergleich von allen 4 Prototypen). Der p-Wert für *Prototyp ohne Fortschrittsanzeige vs. alle anderen mit Fortschrittsanzeige* ist **0.48883** (siehe Anhang 2, S. 8). Diese beiden Werte sind ebenfalls entfernt von einem statistisch signifikanten Wert (**0.05**).

Im Kapitel 5.1 *Definition der User Experience* haben wir die verschiedenen Adjektive den vier Hypothesen zugewiesen.

Die Adjektive *langsam - schnell* passen zu Hypothese 1. Die Auswertung zwischen allen Prototypen wird via Omnibus-Test vorgenommen. Zwischen den Prototypen zeigt sich kein signifikanter Unterschied (p-Wert: **0.128**, siehe Anhang 2, S. 9).

**Somit konnte kein Einfluss durch die Präsenz oder Art einer Fortschrittsanzeige auf die empfundene Ladedauer gezeigt werden.**

Im Kontext dieser beiden Fragen ist der einzige Unterschied zwischen Prototyp 1 (ohne Anzeige) und den anderen Prototypen (mit Anzeige), dass es nur bei Prototyp 1 einen statistisch signifikanten Zusammenhang zwischen der geschätzten Dauer und der empfundenen Angemessenheit gibt. Je länger die Teilnehmer des Prototypen 1 die Dauer schätzten, desto unzufriedener sind sie mit dem Prozess (Spearman Rangkorrelationskoeffizient, der Eindruck der Angemessenheit sinkt im Mittel um 0.29 Skala-Punkte pro Sekunde, siehe Anhang 2, S. 5).

### 7.3.2

## Die visuelle Ästhetik ist bei spielerischer Fortschrittsanzeige am besten

Für die 2. Hypothese nehmen wir an: **«die Zufriedenheit mit der visuellen Ästhetik der Ladeanzeige ist beim Prototypen 4 am grössten»** (siehe Kapitel 5.3).

Vorgängig ist zu erwähnen, dass Prototyp 1 (ohne Fortschrittsanzeige) bei dieser Auswertung nicht berücksichtigt wird, da verschiedene Typen von Fortschrittsanzeigen miteinander verglichen werden.

Folgende Adjektive sind für die Auswertung der Zufriedenheit der visuellen Ästhetik der Ladeanzeige berücksichtigt worden (siehe Kapitel 5.1):

- Unattraktiv – attraktiv,
- Konventionell – originell,
- Unsympathisch – sympathisch,
- Langweilig – spannend

Im Weiteren wollen wir mit der Frage *Empfinden Sie die Anzeige als optisch attraktiv?* herausfinden, welcher der übrigen drei Prototypen die attraktivste Anzeige hat.

Diese Auswertung ist mit dem Chi-Quadrat-Test durchgeführt (Omnibus vergleicht alle Prototypen zusammen). Da der Test signifikant ist, wird ein Posthoc-Test angewendet. Für das Korrekturfverfahren auf der Grundlage des multiplen Testens wenden wir die Bonferroni-Holms Prozedur an.

Die Ergebnisse sind folgendermassen:

- Adjektiv: Unattraktiv – attraktiv  
Prototyp 4 ist attraktiver als Prototyp 2 und 3 (siehe Anhang 2, S. 11).

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00048
- Prototyp 3: korrigierter p-Wert 0.00114

Prototyp 4 hat den höchsten Anteil bei Kategorie 6 (29%) (siehe Anhang 2, S. 12, 2-Way Summary Tabelle).

- Adjektiv: Konventionell – originell  
Prototyp 4 ist origineller als Prototyp 2 und 3 (siehe Anhang 2, S. 13)

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00000
- Prototyp 3: korrigierter p-Wert 0.00000

Prototyp 4 hat den höchsten Anteil bei Kategorie 6 (34%) in der 2-Way Summary Tabelle (siehe Anhang 2, S. 14).

- Adjektiv: Unsympathisch – sympathisch  
Prototyp 4 ist sympathischer als Prototyp 2 und 3 (siehe Anhang 2, S. 15)

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00009
- Prototyp 3: korrigierter p-Wert 0.00001

Prototyp 4 hat den höchsten Anteil bei Kategorie 6 (38%) (siehe Anhang 2, S. 16, 2-Way Summary Tabelle).

- Adjektiv: Langweilig – spannend  
Prototyp 4 ist spannender als Prototyp 2 und 3 (siehe Anhang 2, S. 17)

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00000
- Prototyp 3: korrigierter p-Wert 0.00082

Prototyp 4 hat den höchsten Anteil bei Kategorie 4 (35%) (siehe Anhang 2, S. 18, 2-Way Summary Tabelle).

- Empfinden Sie die Anzeige als optisch attraktiv?

Prototyp 4 ist optisch attraktiver als Prototyp 2 und 3 (siehe Anhang 2, S. 19)

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00000
- Prototyp 3: korrigierter p-Wert 0.00266

Prototyp 4 hat den höchsten Anteil bei Kategorie 4 (53%) (siehe Anhang 2, S. 20, 2-Way Summary Tabelle).

Zusammenfassend sind alle Werte für den Prototypen 4 im Vergleich zu den anderen beiden Prototypen höher. Eine nähere Betrachtung am Beispiel von *Unattraktiv – attraktiv* zeigt, dass gemäss grafischer Darstellung (siehe Anhang 2, S. 12) Prototyp 4 den höchsten Anteil bei Kategorie 6 hat. Im Weiteren zeigt die 2-Way Summary Tabelle den Wert (29%) und insgesamt eine deutliche Tendenz zu höheren Werten. Auch der Pearson Chi-Quadrat-Test (siehe Anhang 2, S. 11) zeigt, dass die Fortschrittsanzeige von Prototyp 4 signifikant attraktiver ist als diejenige der Prototypen 2 und 3.

**Daraus folgt die Schlussfolgerung, dass die Fortschrittsanzeige in Prototyp 4 (mit dem fahrenden Zug) die attraktivste aller getesteten Fortschrittsanzeigen ist. Dasselbe (mit leicht anderen Werten) trifft auf die anderen Adjektive/Variablen zu. Prototyp 4 schneidet hier am besten ab, mit eindeutiger statistischer Signifikanz.**

Die Resultate decken sich somit zu 100% mit Hypothese 2. Der G-Power für unsere Stichprobe von 203 (ohne Prototyp 1 für diese Hypothese) ist gross genug und die Resultate sind somit repräsentativ (siehe G-Power-Kalkulation für die Stichprobe im Anhang 3).

### 7.3.3

## Schlechteste Benutzerzufriedenheit bei Ladedauer ohne Fortschrittsanzeige

Die 3. Hypothese ist: «**Die Zufriedenheit der Ladedauer ist beim Prototyp 3 am grössten**» (siehe Kapitel 5.3).

Um in diesem Fall zu prüfen, ob diese Hypothese valide ist oder ausgeschlossen werden sollte, müssen wir die signifikanten Unterschiede zwischen den 4 Prototypen bezüglich der Frage: *Wie zufrieden sind die Benutzer mit der Ladedauer* herausfinden.

Um die Frage zu beantworten, berücksichtigen wir folgende Adjektive zur Beschreibung (wie auch in Hypothese 1 und 2), wie die Teilnehmer die Anzeige empfunden haben (siehe Kapitel 5.1):

- Nützlich – Nutzlos
- Übersichtlich – Verwirrend
- Angenehm – Lästig

Wie bei den ersten beiden Hypothesen wenden wir den Chi-Quadrat-Test (Omnibus) an, um die Gruppen miteinander zu vergleichen. Der Posthoc-Test wird als zweiter Test angewendet und für das Korrekturverfahren des mehrmaligen Testens der p-Werte wenden wir die Bonferroni-Holms Prozedur an.

Zusammenfassend sind dies die Ergebnisse für die gewählten Adjektive:

- Adjektiv: Nutzlos – nützlich

Der Prototyp 1 (ohne Fortschrittsanzeige) ist weniger nützlich als Prototyp 2 und 3 (siehe Anhang 2, S. 21).

Prototyp 1 vs:

- Prototyp 2: korrigierter p-Wert 0.0327
- Prototyp 3: korrigierter p-Wert 0.0453

- Adjektiv: Verwirrend – übersichtlich

Prototyp 1 (ohne Fortschrittsanzeige) ist verwirrender als alle Prototypen mit Fortschrittsanzeige. Der Prototyp 1 (ohne Fortschrittsanzeige) verhält sich wie folgt mit den anderen Prototypen (siehe Anhang 2, S. 23).

Prototyp 1 vs:

- Prototyp 2: korrigierter p-Wert 0.00000
- Prototyp 3: korrigierter p-Wert 0.00056
- Prototyp 4: korrigierter p-Wert 0.00015

- Adjektiv: Lästig – angenehm

Prototyp 4 ist angenehmer als Prototyp 1 und 2 (siehe Anhang 2, S. 25).

Prototyp 4 vs:

- Prototyp 2: korrigierter p-Wert 0.00505
- Prototyp 1: korrigierter p-Wert 0.00036

**Das Experiment zeigt, dass die Benutzerzufriedenheit bezüglich der Ladedauer für Prototyp 1 (ohne Fortschrittsanzeige) deutlich am schlechtesten ist.**

Für die Aussage *Die Anzeige des Ladevorgangs half, meine Wartezeit kürzer zu machen* wird der Prototyp 1 von der Auswertung ausgeschlossen, da er keine Anzeige hat. Für die restlichen Prototypen ist der korrigierte p-Wert 0.35962 (Anhang 2, Seite 27) deutlich entfernt von dem Signifikanzniveau von 0.05. Es gibt also keinen signifikanten Unterschied zwischen der Anzeige der 3 Prototypen in Bezug auf diese Frage.

#### 7.3.4

### **Die Aussagekraft ist für Prototyp 1 am schwächsten, die Verständlichkeit und Voraussagbarkeit ist bei Prototyp 2 am stärksten**

Bei unserer 4. und letzten Hypothese nehmen wir an: «**Die Aussagekraft der Ladeanzeige ist bei Prototyp 3 am stärksten und bei Prototyp 1 am schwächsten**» (siehe Kapitel 5.3).

Wie bei allen vorgängigen Hypothesen wenden wir für die Auswertung folgende Adjektive an:

- Unberechenbar – voraussagbar
- Verständlich – unverständlich

Auch bei dieser 4. Hypothese wenden wir den Chi-Quadrat-Test (Omnibus) an, um die Gruppen miteinander zu vergleichen. Der Pearson Chi-Quadrat Posthoc-Test wird als zweiter Test angewendet und für das Korrekturverfahren des mehrmaligen Testens der p-Werte wenden wir die Bonferroni-Holms Prozedur an. Weil beim Vergleich von Prototyp 2 und 3 der korrigierte p-Wert nur ganz knapp über der 0.05 Grenze liegt, wenden wir einen zusätzlichen Test, Maximum-Likelihood (ML) Chi-Quadrat-Test, an, um dieses Resultat zu überprüfen. Beim ML Chi-Quadrat-Test ist der korrigierte p-Wert unter der 0.05-Schwelle und damit signifikant.

- Unberechenbar – voraussagbar

Prototyp 2 ist im Vergleich voraussagbarer als alle anderen Prototypen (siehe Anhang 2, S. 28).

Prototyp 2 vs:

- Prototyp 1: korrigierter p-Wert 0.00000
- Prototyp 3: korrigierter p-Wert 0.05667 (Pearson Chi-Quadrat), 0.01611 (ML Chi-Quadrat)
- Prototyp 4: korrigierter p-Wert 0.0072

Darüber hinaus ist Prototyp 1 unberechenbarer als Prototyp 3, korrigierter p-Wert: 0,03568 (siehe Anhang 2, S. 28)

- Verständlich – unverständlich

Prototyp 1 ist unverständlich, mit sehr klarem Unterschied zu allen anderen Prototypen (siehe Anhang 2, S. 30).

Der korrigierte p-Wert: 0.00000 steht im Vergleich zu Prototyp 2, 3 und 4.

Darüber hinaus ist Prototyp 2 verständlicher als Prototyp 3 und 4 (siehe Anhang 2, S. 30).

Prototyp 2 vs:

- Prototyp 3: korrigierter p-Wert 0.0365
- Prototyp 4: korrigierter p-Wert 0.03976

Der G-Power für unsere Stichprobe von 255 ist gross genug und somit sind die Resultate repräsentativ (Siehe G-Power-Kalkulation für die Stichprobe im Anhang 3).

**Wir haben jedoch die Hypothese für Prototyp 3 nicht validiert. Wir können nicht bestätigen, dass die Aussagekraft von Prototyp 3 am stärksten ist.**

**Prototyp 2 jedoch entpuppt sich als Gewinner, da wir validieren konnten, dass die Fortschrittsanzeige verständlicher und voraussagbarer ist als die aller anderen Prototypen.**





## 8. Diskussion

Wie im Kapitel 7 beschrieben, konnten wir beweisen, dass ein Prototyp ohne Feedback keinen Gefallen findet und die Wartezeit von den Benutzern als unangemessen eingestuft wird. Die Befragten werteten den Prototypen ohne Feedback als am wenigsten aussagekräftig. Diese Resultate bestätigen die Aussage von mehreren Studien (Johnson, 2008; Nielsen, 1993; Card et al., 1991; Seow, 2008), dass die Benutzer Feedback benötigen, auch wenn die wahrgenommene Wartezeit durch das Feedback verlängert wird (Branaghan/Sanchez, 2009: 528f.). Wir konnten nicht statistisch beweisen, dass ein Feedback hilft, die wahrgenommene Wartezeit als kürzer zu empfinden oder dass der Benutzer die Wartezeit als begründet wahrnimmt. Dazu waren die Testresultate nicht signifikant genug.

Auch konnten wir nicht statistisch signifikant nachweisen, dass die gemäss Literaturrecherche (Seow, 2008; Mishunov, 2015c) ideale Ladeanzeige für eine Wartezeit von 7 Sekunden (Prototyp 3, Fortschrittsanzeige mit Text) von den Benutzern als passend eingeschätzt wurde. Auch stufen die Befragten die Wartezeit weder als angemessen noch als benutzerfreundlicher ein.

Wir konnten aber beweisen, dass die befragten Benutzer die Anzeige im Prototypen 4 als attraktiver und sympathischer einstufen als die anderen Anzeigen, obwohl es eine *Busy*-Anzeige ist. Gemäss Literatur (2008: 80f.) passt diese Anzeige nicht zur Wartezeit von 7 Sekunden. Wir vermuten, dass die Benutzer wegen der spielerischen Gestaltung die Wartezeit nicht als negativ wahrnehmen. An diesem Punkt könnte man die Forschung fortsetzen und untersuchen, ob spielerisch gestaltete Indikatoren einen positiven emotionalen Einfluss auf die Wartezeit im Web haben.

Eine weitere Erkenntnis aus unserer Studie ist, dass der Fortschrittsbalken für die Benutzer die vertrauteste Anzeigeform ist. Die Benutzer stufen den Fortschrittsbalken als voraussagbarer und verständlicher als alle anderen Anzeigen ein. Der Grund dafür könnte die weite Verbreitung dieser Anzeige sein. So empfiehlt auch Google in seinen Material Design Guidelines die Benutzung des Fortschrittsbalken unter anderem für ladende Webseiten. Benutzer sind daher an diese Anzeige gewöhnt und mit ihr vertraut.

Es gibt jedoch mehrere Faktoren, die bei der Bewertung der Ergebnisse der vorliegenden Forschung berücksichtigt werden müssen.

In Anbetracht unserer Resultate werfen wir einen selbstkritischen Blick auf die Skala im Fragebogen. Man kann in den Antworten eine Tendenz zum mittleren Wert der benutzten Likert-Skala (siehe Kapitel 7.2 *Deskriptive Daten*) feststellen. Das kann unserer Meinung nach zwei Gründe haben. Einerseits verlangt die Frage mit den gegensätzlichen Adjektiven eine ziemlich konkrete Stellungnahme von unseren Befragten, und auch wenn der Befragte vielleicht unsicher war, musste er die Frage beantworten. Dies kann dazu führen, dass bei gewissen Fragen im Zweifelsfalle der Mittelwert gewählt wurde. Andererseits ist bekannt, dass Personen in Umfragen zu eher neutralen Antworten tendieren. Es geht dabei um soziale Anpassung, man will sich mit extremen Meinungen nicht exponieren (Courage/Baxter, 2005: 315). Deshalb denken wir, dass wir besser eine 6er-Skala statt einer 5er- und 7er-Skala hätten verwenden sollen. So wären die Befragten mit der geraden Anzahl von möglichen Antworten zu einer klaren Aussage gezwungen worden und die Unterschiede in der Bewertung wären deutlicher zum Vorschein gekommen. Mit zusätzlichem

Aufwand hätten die Daten gruppiert ausgewertet werden können, denn durch Gruppenbildung hätten die Extreme besser berücksichtigt werden können. Eine erneute Auswertung des Datensatzes wäre aber sehr aufwändig gewesen, weshalb wir uns dagegen entschieden haben.

Weiter ist zu berücksichtigen, dass eine Umfrage mit einem Prototypen trotz grosser Detailtreue unter künstlichen Bedingungen stattfindet. Wer wirklich ein Billett für eine anstehende Reise benötigt und dies womöglich unter Zeitdruck erstehen will, reagiert ganz anders als während einer Umfrage, in der der Billettkauf nur fiktiv ist. Dies kann dann auch einen Einfluss auf die Antworten haben, die so vielleicht weniger prägnant ausfallen. Eine spätere Beobachtung des Verhaltens der Benutzer unter realen Bedingungen im Ticketshop könnte hier zusätzliche Information liefern.

## 8.1

### Empfehlung an die SBB

Unsere konkrete Empfehlung an die SBB ist, Wartezeiten über 2 Sekunden im Ticket-Shop mit einer Fortschrittsanzeige zu überbrücken. Wir empfehlen, diese Fortschrittsanzeige auf dem Warenkorb zu positionieren, gemäss unseren Usability-Kenntnissen, Konkurrenzanalysen und entsprechend der Literaturrecherche (Wroblewski, 2014; Zumbrennen, 2015; Heal, 2016: 116).

Die Befragung hat ergeben, dass sich der Fortschrittsbalken am besten eignet (Abbildung 8.1). Die Befragten werten diesen als verständlich und voraussagbar, solange die Wartezeit bekannt ist. Wenn die Wartezeit nicht berechnet werden kann, empfehlen wir trotz fehlender Signifikanz in den Resultaten die Benutzung einer *Class C*-Anzeige (Abbildung 8.2), die während des Wartens zusätzliche Informationen über den Prozess liefert (Seow, 2008: 80f.; Mishunov, 2015c). Diese Anzeige lässt sich mit einer horizontalen Anzeige (ähnlich zu einem Fortschrittsbalken, siehe *Google Material Design*) kombinieren. Da unsere Befragten die spielerische Anzeige mit dem Zug als optisch attraktiv eingeschätzt haben, kann auch diese spielerische Variante mit einer horizontalen Anzeige und einem zum Thema passenden Indikator (z.B. einem Zug, der durch den Warenkorb fährt, analog zu unserem Prototyp 4) als Alternative eingesetzt werden (Abbildung 8.3).

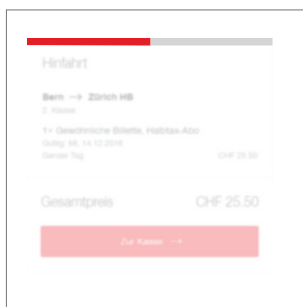


Abbildung 8.1: Fortschrittsbalken

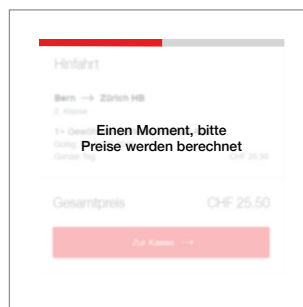


Abbildung 8.2: Horizontale Anzeige + Text

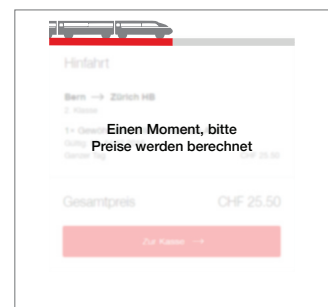


Abbildung 8.3: Zug-Balken + Text

## 8.2

### Exkurs: Expert-Review Ergebnisse

Als UX-Designer lassen wir es uns an dieser Stelle nicht entgehen, einen Expert-Review zur neuen Shop-Seite zu erstellen, da uns ein paar Verbesserungsmöglichkeiten ins Auge sprangen.

Zu folgenden Elementen der Reiseoptionenseite haben wir Verbesserungsvorschläge:

Im Allgemeinen verzichtet man auf ein Dropdown, wenn weniger als sieben Optionen angezeigt werden. Alternative Möglichkeiten dafür wären: für zwei Optionen einen Radiobutton darzustellen oder einen Toggle Button zu verwenden (Abbildung 8.4).

Das Label im Eingabefeld wird nach «Material Design Style Guide» direkt im Eingabefeld (In-line Label) dargestellt. Das wurde bei der Mobile App der SBB bereits so umgesetzt. Um den Benutzer nicht zu verwirren und die Vertrautheit zu behalten, empfehlen wir, dies in gleicher Art umzusetzen (Abbildungen 8.5).

Die Informationen im Warenkorb sind unübersichtlich dargestellt. Im Usability-Test haben wir herausgefunden, dass der Begriff *Gewöhnliche Billette* für den Benutzer unklar ist. Man könnte stattdessen schreiben *Billette ohne Ermässigung*. Auch geht die Anzahl der Billette in dieser Darstellung unter. Die Anzahl könnte über den Preis angeordnet werden (Abbildung 8.6).

The image shows a user interface for selecting a discount. Under the heading "Aktuelle Lösung", there is a dropdown menu labeled "Ermässigungskarten" with "Halbtax" selected. Below this, under "Unsere Empfehlung", there are four radio button options: "Keine Ermässigung", "Halbtax Abo", "Junior-Karte", and "6-16 Jahre".

Abbildung 8.4: Ermässigungs-Dropdown der SBB und unsere Empfehlung

The image shows two input fields side-by-side. The left one is labeled "Aktuelle Lösung" and "Vorname" with an empty input box. The right one is labeled "Unsere Empfehlung" and "Vorname" with the text "Hello" inside the input box.

Abbildung 8.5: Eingabefeld SBB und unsere Empfehlung

The image shows a shopping cart item for a one-way ticket. It includes the text "Hinfahrt", "Bern -> Zürich HB", "2. Klasse", "1 x Gewöhnliche Billette", "Halbtax-Abo", "Gültig: Mi, 14.12.2016", "Ganzer Tag", and "CHF 25.50". Below this, it shows "Gesamtpreis CHF 25.50" and a red button labeled "Zur Kasse ->".

Abbildung 8.6: Wording auf dem Warenkorb



## 9. Reflexion

Als wir uns für dieses Masterarbeitsthema entschieden wussten wir, dass es eine anspruchsvolle Arbeit wird, für die uns vor allem relevantes Wissen in Psychologie und Statistik fehlt. Es hat sich tatsächlich auch gezeigt, dass es zwar schwierig war, aber nichtsdestotrotz machbar. Nebst dem intensiven Studium der nötigen Literatur hat uns vor allem unser Netzwerk geholfen, durch das wir über Freunde und Bekannte schnell zu relevanten Informationen kamen und uns so auf die wichtigen Teile der neuen Wissensgebiete fokussieren konnten.

Wir alle erstellten zum ersten Mal eine wissenschaftliche Arbeit, was uns mehrmals fast zur Verzweiflung brachte, da es oft nur einen richtigen Weg gibt und Fehler im Vorgehen später nur schwer auszubügeln sind. Hier waren es die gegenseitige Unterstützung im Team und die Motivation aus unserer Umgebung, die uns immer wieder Mut machten, weiterzumachen.

Zu Beginn der Arbeit dachten wir, dass wir den statistischen Teil selber auswerten können, hatten aber dann bemerkt, dass uns dies viel Zeit kostet, vor allem, sich das Wissen im Umgang mit den Tools zu erarbeiten, um die die grosse Datenmenge richtig auswerten zu können. Wir hatten daher nach Rücksprache mit unserem Betreuer eine Expertin in unserem Netzwerk um Unterstützung gebeten, die die statistische Auswertung übernahm. Wir haben viel gelernt über Statistik: von der Planung zur Umsetzung bis zur Auswertung und Interpretation der Daten.

Gerade beim Anwenden einer quantitativen Methode haben wir ebenfalls viel gelernt. In der Praxis wird sehr häufig mit qualitativen Methoden gearbeitet. Es war sehr lehrreich zu sehen, was es braucht, bis eine Aussage wirklich statistische

Relevanz erhält. Wir haben auch gesehen, wie aufwändig quantitative Methoden sind, weshalb sie in der Praxis nur zurückhaltend eingesetzt werden. Wir haben in dieser Arbeit auch die Effektivität in der Kombination von quantitativen und qualitativen Methoden erfahren. Mit dem Usability-Testing konnten wir unseren Fragebogen wirksam weiterentwickeln und für den praktischen Einsatz austesten.

Die Entwicklung des Fragebogens war der schwierigste Teil der Arbeit, und wir unterschätzten ihn zu Beginn. Auch haben wir in dieser Phase den Fehler gemacht, Skalen mit ungeraden Antwortmöglichkeiten zu wählen, was wir heute nicht mehr so tun würden. Der Einsatz von Standardfragebogen hätte uns zwar viel Kopfzerbrechen erspart, wir sind aber nach wie vor überzeugt, dass wir mit unserem selbst entwickelten Fragebogen spezifischer auf unsere Studie eingehen konnten und die Resultate präziser waren und unserer Auftraggeberin einen grossen Mehrwert bieten, auch wenn die statistische Signifikanz teilweise fehlt.

Als Team haben wir sehr gut funktioniert, auch wenn wir uns bei vielen Punkten oft nicht einig waren, fanden wir doch immer eine Lösung und konnten gemeinsam eine für alle anspruchsvolle Arbeit erfolgreich beenden.

Sehr positiv war die Unterstützung aus unserem Umfeld, sei es mit Tipps, Know-How oder auch der Tatsache, dass viele an der Umfrage teilnahmen. Ohne die 282 Personen hätten wir diese Arbeit nicht erfolgreich abschliessen können.

# 10. Literaturverzeichnis

Akamai Technologies (2013): Innovating at the Pace of Mobile: Akamai Aqua Ion Whitepaper [Website], Slideshare. Online unter: <http://www.slideshare.net/AkamaiTechnologies/akamai-aqua-ionwhitepaper> (Zugriff am 20.12.16).

Akamai Technologies. (2009): Mobile Web User Expectations: 2009 vs 2011 [Pressemitteilung], online unter: <https://www.akamai.com/de/de/about/news/press/2009-press/akamai-reveals-2-seconds-as-the-new-threshold-of-acceptability-for-ecommerce-web-page-response-times.jsp> (Zugriff am 2.11.16).

Albert, W./ Tullis, T. (2013): Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics. 2. Auflage. San Francisco: Morgan Kaufmann

Belshe, M. (2010): More Bandwidth Doesn't Matter (Much) [Blogeintrag], online unter: <https://www.belshe.com/2010/05/24/more-bandwidth-doesnt-matter-much/> (Zugriff am 23.6.16)

Bixby, J. (2011): Web Performance Today [Website], Web Performance Today. Online unter: <http://www.webperformancetoday.com/2011/07/20/new-findings-mobile-web-users-are-more-disappointed-than-ever/> (Zugriff am 7.7.16).

Branaghan, R. J./ Sanchez, C.A. (2009): Feedback Preferences and Impressions of Waiting, in *Human Factors*, Vol. 51, No. 4, S. 528-538

Brooke, J. (1996): SUS: a "quick and dirty" usability scale. In: Jordan, P. W./ Thomas B./ Weerdmeester, B. A./ McClelland, A. L.: *Usability Evaluation in Industry*. London: Taylor and Francis, S. 107-114.

Brown, D. M. (2007): Communicating Design: Developing Web Site Documentation for Design and Planning. 1. Auflage. Berkeley: New Riders

Card, S. K./ Robertson, G. G./ Mackinlay, J. D. (1991): The information visualizer: An information workspace [CHI '91 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, Mai 1991], S. 181-188.

Coolican, H. (2014): Research Methods and Statistics in Psychology. 6. Auflage. London [u.a.]: Psychology Press

Courage, C./ Baxter K. (2005): Understanding your users. A practical guide to user requirements. Methods, Tools, & Techniques. 1. Auflage. San Francisco: Morgan Kaufmann

Doherty, R./ Sorensen, P. (2015): Keeping users in the flow: Mapping system responsiveness with user experience, in *Procedia Manufacturing*, Volume 3, 2015, S. 4384-4391

- Field, A./ Hold, G. (2010): How to design and report experiments. Reprint. London [u.a.]: SAGE
- Fui-Hoon Nah, F. (2004): A study on tolerable waiting time: how long are Web users willing to wait? in *Behavior and Information Technology*, Vol. 23, No. 3, S. 153-163
- Goodwin, D. (2009): Designing for the Digital Age. 1. Auflage. Indianapolis: Wiley
- Google Inc. (2014): Material Design. Components – Progress & activity, online unter: <https://material.io/guidelines/components/progress-activity.html> (Zugriff am 6.1.17)
- Gulliksen, J./ Göransson, B./ Boivie, I./ Blomkvist, S./ Persson, J./ Cajander, Å. (2003): Key principles for user-centred systems design, in *Behaviour & Information Technology*, Vol. 22, No. 6, S. 397–409
- Harrison, C./ Amento, B./ Kuznetsov, S./ Bell, R. (2007): Rethinking the Progress Bar [UIST '07 Proceedings of the 20th annual ACM symposium on User interface software and technology, Newport, Oktober 2007], S. 115-118.
- Head, V. (2016): Designing Interface Animation: Meaningful Motion for User Experience. 1. Auflage. New York: Rosenfeld Media
- Hohenstein, J./ Khan, H./ Canfield, K./ Tung, S./ Perez Cano, R. (2016): Shorter Wait Times: The Effects of Various Loading Screens on Perceived Performance [CHI EA '16 Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Santa Clara, Mai 2016], S. 3084-3090
- Hübscher, C. (2015): Vorgehensmodelle User Centered Design II. Skript Modul Vorgehensmodelle II 2015/16. Rapperswil: Hochschule für Technik Rapperswil
- Johnson, J. (2008): GUI Bloopers 2.0. Common User Interface Design Don'ts and Do. 2. Auflage. Burlington: Morgan Kaufmann
- Johnson, J. (2010): Design with the mind in mind, 2. Auflage. Waltham: Morgan Kaufmann
- King, A. (2008): The Psychology of Web Performance [Website], Website Optimization. Online unter: <http://www.websiteoptimization.com/speed/tweak/psychology-web-performance/> [Accessed 2. Dec 2016].
- Klein S./ Porst, R. (2000): Mail Surveys. Ein Literaturbericht, online unter: [http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2000/00\\_10.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2000/00_10.pdf) (Zugriff am 12.1.17)

Lallemand, C./ Gronier, G. (2012): Enhancing User eXperience during Waiting Time in HCI: Contributions of Cognitive Psychology. [Proceedings of the Designing Interactive Systems Conference, Newcastle Upon Tyne, Juni 2012], S. 751-760

Laugwitz, B./ Held, T./ Schrepp, M. (2008): Construction and Evaluation of a User Experience Questionnaire, [USAB '08 Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work, Graz, November 2008], S. 63-76

Lederer, Bernd (2015): Quantitative Erhebungsmethoden. In: Hug, T./ Poscheschnik, G. (2015): *Empirisch forschen. Die Planung und Umsetzung von Projekten im Studium*. 2., überarbeitete Auflage. Konstanz: UTB, S. 110-133

Mair-Noack, M. (2007): Billettautomaten: Die Technik überfordert viele [Artikel], online unter: <https://www.ktipp.ch/a1029936> (Zugriff am 29.9.16)

McInerney, P./ Li, Jin (2002): Progress Indication: Concepts, Design, and Implementation [Artikel], online unter: <http://www.ibm.com/developerworks/web/library/us-progind/> (Zugriff am 31.7.16)

Miller, R. B. (1968): Response time in user-system conversational transactions [Proceedings of the AFIPS Fall Joint Computer Conference Part I, San Francisco, Dezember 1968], S. 267-277.

Minitab 17 (2016): What is randomization? [Website], online unter: <http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/doe/basics/what-is-randomization/> (Zugriff am 10.1.16)

Mishunov, D. (2015a): Why Performance Matters, Part 1: The Perception of Time [Website], Smashing Magazine. Online unter: <https://www.smashingmagazine.com/2015/09/why-performance-matters-the-perception-of-time/> (Zugriff am 20.11.16)

Mishunov, D. (2015b): Why Performance Matters, Part 2: Perception Management [Website], Smashing Magazine. Online unter: <https://www.smashingmagazine.com/2015/11/why-performance-matters-part-2-perception-management/> (Zugriff am 20.11.16).

Mishunov, D. (2015c): Why Performance Matters, Part 3: Tolerance Management [Website], Smashing Magazine. Online unter: <https://www.smashingmagazine.com/2015/12/performance-matters-part-3-tolerance-management/> (Zugriff am 20.11.16).

Myers, B. A. (1985): The importance of percent-done progress indicators for computer-human interfaces [CHI '85 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Francisco, April 1985], S. 11-17.

Nielsen, J. (1993): Usability Engineering. 1. Auflage. San Diego: Academic Press

Nielsen, J. (1995): Severity Ratings for Usability Problems [Artikel], online unter: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/> (Zugriff am 13.1.16)

Nielsen, J. (1996): International Usability Testing [Artikel], online unter: <https://www.nngroup.com/articles/international-usability-testing/> (Zugriff am 20.12.2016)

Norman, D. (2004): Emotional Design: Why We Love (Or Hate) Everyday Things. 1. Auflage. New York: Basic Books

Reiss, S./ Sarris, V. (2012): Experimentelle Psychologie, 1. Auflage. München: Pearson

Ronan, C. (2016): The Web is Doom [Website], Mobiforge. Online unter: <https://mobiforge.com/research-analysis/the-web-is-doom> (Zugriff am 23.12.16).

Sarodnick, F./ Brau, H. (2011): Methoden der Usability Evaluation. 1. Auflage. Bern: Verlag Hans Huber

Sauro, J. (2011): Measuring Usability with the System Usability Scale (SUS) [Blogeintrag], online unter: <http://www.measuringu.com/sus.php> (Zugriff am 18.12.16)

Sauro, J./ Lewis, J. R. (2016): Quantifying the User Experience. Practical Statistics for User Research. 2. Auflage. Cambridge: Morgan Kaufmann

Schrepp, M. (2015): User Experience Questionnaire Handbook. All you need to know to apply the UEQ successfully in your projects, online unter: <http://www.ueq-online.org/ueq-download/> (Zugriff am 16.10.16)

Schurman, E./ Brutlag, J. (2009): The User and Business Impact of Server Delays [Website], Slideshare. Online unter: <http://www.slideshare.net/dyninc/the-user-and-business-impact-of-server-delays-additional-bytes-and-http-chunking-in-web-search-presentation> (Zugriff am 19.12.16).

Seow, S. C. (2008): Designing and Engineering Time: The Psychology of Time Perception in Software. 1. Auflage Boston: Addison-Wesley

Souders, Steve. (2009): Velocity and the Bottom Line [Blogeintrag], online unter: <http://radar.oreilly.com/2009/07/velocity-making-your-site-fast.html> (Zugriff am 23.9.16).

Thielsch, M. T./ Jaron, R. (2012): Das Zusammenspiel von Website-Inhalten, Usability und Ästhetik. In: Reiterer, H./ Deussen, O. (Hrsg.): *Mensch & Computer 2012: interaktiv informiert – allgegenwärtig und allumfassend!?* München: Oldenbourg Verlag, S. 123-132.

Wharton, C./ Rieman, J./ Lewis, C./ Polson, P. (1994): The Cognitive Walkthrough Method: A Practitioner's Guide. In: Nielsen, J./ Mack, R. L.: *Usability Inspection Methods*, New York: John Wiley & Sons, S. 105 - 140.

Wroblewski, L. (2013): Mobile Design Details: Avoid The Spinner [Blogeintrag], online unter: <http://www.lukew.com/ff/entry.asp?1797> (Zugriff am 4.12.16)

Wroblewski, L. (2014): Video: How to Avoid Loading Indicators [Video], online unter: <http://www.lukew.com/ff/entry.asp?1904> (Zugriff am 25.7.16)

Zumbrunnen, A. (2015): The Illusion of Time [Website], Medium. Online unter: <https://medium.com/swlh/the-illusion-of-time-8f321fa2f191> (Zugriff am 4.12.16)



# 11. Abbildungsverzeichnis

<b>Abbildung 1.1:</b> Aktuelle Webshop der SBB	14	<b>Abbildung 5.2:</b> Durschnittsladedauer zwischen der Seite Reisende und der Seite Reiseangebote	45
<b>Abbildung 1.2:</b> Wireframe der Reiseoptionenseite (von uns gemacht)	15	<b>Abbildung 5.3:</b> Aktuelle Fortschrittsanzeige der SBB	45
<b>Abbildung 2.1:</b> Typen von Zeitspannen (Seow, 2008)	20	<b>Abbildung 5.4:</b> Fortschrittsanzeigen der Studie	46
<b>Abbildung 2.2:</b> Übersicht von Fortschrittsanzeigen (Seow, 2008)	21	<b>Abbildung 5.5:</b> Tabelle: Variable der Studie	47
<b>Abbildung 2.3:</b> Beispiele von Spinners	22	<b>Abbildung 5.6:</b> Tabelle: Unsere Prototypen	51
<b>Abbildung 2.4:</b> Fortschrittsanzeige Outlook	23	<b>Abbildung 5.7:</b> Fortschrittsanzeige mit dem Zug	51
<b>Abbildung 3.1:</b> Mike Morgenthaler	28	<b>Abbildung 6.1:</b> Fahrplanansicht aus dem alten Prototyp	54
<b>Abbildung 3.2:</b> Backend der SBB Webshop	29	<b>Abbildung 6.2:</b> Alter und neuer Spinner	58
<b>Abbildung 3.3:</b> Tabelle: Resultate der Konkurrenzanalyse (Brown, 2007)	33	<b>Abbildung 6.3:</b> Fortschrittsbalken mit Prozenten	59
<b>Abbildung 4.1:</b> Wissenschaftliche Methode (Reiß/Sarris, 2012)	35	<b>Abbildung 6.4:</b> Fotos aus dem Usability Testing	61
<b>Abbildung 4.2:</b> Tabelle: Checkliste unserer experimentellen Studie (Reiß/Sarris, 2012)	39	<b>Abbildung 6.5:</b> Usability Räumlichkeiten	62
<b>Abbildung 4.3:</b> Vergleich quantitatives und qualitatives Vorgehen (Lederer, 2015)	38	<b>Abbildung 6.6:</b> Storyboard	64
<b>Abbildung 5.1:</b> Durschnittsladedauer zwischen der Seite Online-Fahrplan und der Seite Reisende	45	<b>Abbildung 6.7:</b> Finale Version der Fortschrittsanzeigen	66
		<b>Abbildung 6.8:</b> Workshop-Resultate für die Entwicklung des Fragebogens	67

<b>Abbildung 6.9:</b> Mapping Fragen und Adjektiven zu den Hypothesen	69	<b>Abbildung 8.5:</b> Eingabefelder und unsere Empfehlung	87
<b>Abbildung 6.10:</b> Kontrollfrage im Fragebogen	68	<b>Abbildung 8.6:</b> Wording auf dem Warenkorb	87
<b>Abbildung 6.11:</b> Tabelle: Anzahl versendeter E-Mails pro Tag	71		
<b>Abbildung 7.1:</b> Graphik: Geschlecht und Alter	75		
<b>Abbildung 7.2:</b> Graphik: Zugfahrverhalten	76		
<b>Abbildung 7.3:</b> Graphik: Billettkaufoptionen	76		
<b>Abbildung 7.4:</b> Tabelle: Durschnitt der wahrgenommenen Ladezeit pro Prototyp	75		
<b>Abbildung 7.5:</b> Zentralwert-Berechnung aller Prototypen	77		
<b>Abbildung 7.6:</b> Zentralwert-Berechnung pro Prototyp	77		
<b>Abbildung 8.1:</b> Empfehlung: Fortschrittsbalken	86		
<b>Abbildung 8.2:</b> Empfehlung: Horizontale Anzeige mit Text	86		
<b>Abbildung 8.3:</b> Empfehlung: Zug-Balken mit Text	86		
<b>Abbildung 8.4:</b> Ermässigungs-Dropdown der SBB und unsere Empfehlung	87		

## 12. Glossar

Referenz	Dokument
Behavioural (Designlevel)	Die antrainierte, assoziative Reaktion und Wahrnehmung beim Menschen (Quelle: Don Norman, jnd.org)
Bonferroni-Holms Prozedur	Prozedur um die Korrektur des Fehlers erster Art (Alphafehler-Kumulierung) durchzuführen. Die Alphafehler-Kumulierung bezeichnet in der Statistik die globale Erhöhung der Alpha-Fehler-Wahrscheinlichkeit durch multiples Testen in derselben Stichprobe (Quelle: Wikipedia)
Captive	Zeitspanne für Systemantwortung zwischen 7 und 10 Sekunden (Seow, 2008)
Class A(ctive)	Typ von Fortschrittsanzeige gemäss Seow (2008), beispielsweise ein Fortschrittsbalken
Class B(usy)	Typ von Fortschrittsanzeige gemäss Seow (2008), beispielsweise die drehende Sanduhr von Microsoft Windows
Class C(hanges)	Typ von Fortschrittsanzeige gemäss Seow (2008), beispielsweise die Fortschrittsanzeige eines Antiviren-Programm
Class D(isappear)	Typ von Fortschrittsanzeige gemäss Seow (2008). Benutzer wird direkt über das Ende des Prozesses informiert. Dieser läuft dann aber im Hintergrund ohne weiteres Feedback und lässt den Benutzern so die Freiheit, währenddessen eine andere Tätigkeit zu verrichten, z.B. bei einem langen Download
Continuous	Zeitspanne für Systemantwortung zwischen 2 und 5 Sekunden (Seow, 2008)
CSS	Cascading Stylesheets (CSS) sind Anweisungen an den Browser zur Darstellung einer Webseite (Quelle: Wikipedia)
Datensatz	Im statistischen Sinne ist ein Datensatz die Gesamtheit aller gesammelten Daten, die sich in einzelne Fällen (siehe Fall) unterteilen lässt.

Fall	Im Datensatz, jeder Eintrag, Zeile oder Antwort der Umfrage.
Findings	Ergebnisse aus einem Test, wie z.B. Usability-Test.
Fortschrittsanzeige	Eine optische Anzeige bei Computerprogrammen, die darstellen soll, wie weit die Bearbeitung eines Auftrags fortgeschritten ist bzw. dass die Bearbeitung im Gange ist (Quelle: Wikipedia)
Fortschrittsbalken	siehe Fortschrittsanzeige
G-Power	G-Power ist ein Mittel statistische Poweranalysen für verschiedene Tests durchzuführen: t-Tests, F-Tests, Chi-Quadrat-Tests, z-Tests und einige exakte Tests. G-Power kann auch dazu benutzt werden, verschiedene Effektstärken zu berechnen und zu grafischen Darstellungen verschiedener Poweranalysen dienen (Quelle: Universität Düsseldorf)
GA	Generalabonnement ist ein Ticket der SBB, mit dem man freie Fahrt mit Zügen der SBB sowie Fähren, Trams und Bussen hat (Quelle: SBB.ch)
high-fidelity Prototypen	Prototypen, die eine höhere Ähnlichkeit mit dem endgültigen Produkt haben. Sie werden mit Hilfe von Software entwickelt.
Immediate	Zeitspanne für Systemantwortung zwischen 0.5 und 1 Sekunden (Seow, 2008)
Instantaneous	Zeitspanne für Systemantwortung zwischen 100 und 200 Millisekunden (Seow, 2008)
Kruskal-Wallis-Rangsummentest	Der Kruskal-Wallis-Test für unabhängige Stichproben testet, ob sich die zentralen Tendenzen mehrerer unabhängiger Stichproben unterscheiden (Quelle: Wikipedia)
Ladeanzeige	Siehe Fortschrittsanzeige
Ladeverhalten	Als Ladeverhalten in unserer Arbeit wird das Verhalten eines Systems beim Übergang eines Systemzustandes zum nächsten Systemzustand bezeichnet.

Ladevorgangs- anzeige	Siehe Fortschrittsanzeige
low-fidelity Prototypen	Prototypen, die auf Basis von Skizzen und Zeichnungen auf Papier gemacht werden.
Mann-Whitney- U-Test	Der Mann-Whitney-U-Test für unabhängige Stichproben testet, ob die zentralen Tendenzen zweier unabhängiger Stichproben verschieden sind (Quelle: Universität Zürich)
Omnibus-Test	Der Test prüft nur ob es einen Unterschied zwischen mehreren Gruppen (Grundgesamtheiten) gibt oder nicht. Welche Gruppen für den Unterschied verantwortlich sind, dafür werden im Anschluss Post-Hoc-Tests durchgeführt. Der Test gibt keine Auskunft darüber.
Pearson Chi-Quadrat-Test	Der Pearson Chi-Quadrat-Test testet, ob zwischen zwei kategorialen Variablen ein Zusammenhang besteht (Quelle: Universität Zürich)
Post-hoc-Tests	Post-hoc-Tests sind Signifikanztests aus der mathematischen Statistik. Die Post-hoc-Tests ermöglichen durch gruppenweise Vergleiche eine Aussage darüber, welche Gruppen-Mittelwerte nicht signifikant verschieden sind (Quelle: Wikipedia)
Randomisierung	Verfahren, bei dem die Versuchspersonen (zum Beispiel teilnehmende Patienten) unter Verwendung eines Zufallsmechanismus unterschiedlichen Gruppen zugeordnet werden (Quelle: Wikipedia)
Reflective (Designlevel)	Die gedankliche und verarbeitende Reaktion und Wahrnehmung beim Menschen (Quelle: Don Norman, <a href="http://www.jnd.org/">http://www.jnd.org/</a> )
Responsiveness	Die Fähigkeit der Anwendungssoftware, den Benutzern informiert zu halten, auch wenn die Software einen User Request nicht unmittelbar beantworten kann (Johnson, 2008)
Signifikanz	Statistisch signifikant wird das Ergebnis eines statistischen Tests genannt, wenn Stichprobendaten so stark von einer vorher festgelegten Annahme (der Nullhypothese) abweichen, dass diese Annahme nach einer vorher festgelegten Regel verworfen wird (Quelle: Wikipedia)
Single-Page-App	Bei einer Single-Page-Applikation werden alle notwendigen Ressourcen (Html, CSS, Javascript) entweder auf einmal geladen oder dynamisch nachgeladen. Ein komplettes Neuladen der Seite ist somit nicht notwendig (Quelle: Wikipedia).

Spearmans Rangokorrelationskoeffizient	Es ist ein Typ von Rangkorrelationskoeffizient. Ein Rangkorrelationskoeffizient ist ein parameterfreies Mass für Korrelationen, das heisst, er misst, wie gut eine beliebige monotone Funktion den Zusammenhang zwischen zwei Variablen beschreiben kann, ohne irgendwelche Annahmen über die Wahrscheinlichkeitsverteilung der Variablen zu machen (Quelle: Wikipedia)
Spinner	Ein Typ von Fortschrittsanzeige, die aus einem drehenden Element besteht, z.B. Microsoft Windows' drehende Sanduhr
Spinning Wheel	Siehe Spinner
Stichprobe	Als Stichprobe bezeichnet man eine Teilmenge einer Grundgesamtheit, die unter bestimmten Gesichtspunkten ausgewählt wurde (Quelle: Wikipedia)
Szenario	Ein Szenario beschreibt den Kontext der Interaktion mit einem System. Oftmals ist das Szenario für ausgewählte Benutzergruppen in Form einer Geschichte dargestellt, um die Benutzer emotional stärker einzubinden. Es werden Möglichkeiten, Ziele und Fragen aufgeführt, die den Benutzer bei der Interaktion betreffen (Quelle: Nielsen, <a href="https://www.nngroup.com/articles/task-scenarios-usability-testing/">https://www.nngroup.com/articles/task-scenarios-usability-testing/</a> )
Variable (Statistik)	Eine statistische Variable ist eine messbare oder zählbare Eigenschaft, Zahl oder Menge (Quelle: Australian Bureau of Statistics, <a href="http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+what+are+variables">http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+what+are+variables</a> )
Versuchsplan (Statistik)	Die statistische Versuchsplanung umfasst alle statistischen Verfahren, die vor Versuchsbeginn angewendet werden sollten (Quelle: Wikipedia)
Visceral (Designlevel)	Die unmittelbare, schnelle und elementare Reaktion und Wahrnehmung beim Menschen (Quelle: Don Norman, <a href="http://www.jnd.org/">http://www.jnd.org/</a> )



# 13. Anhang

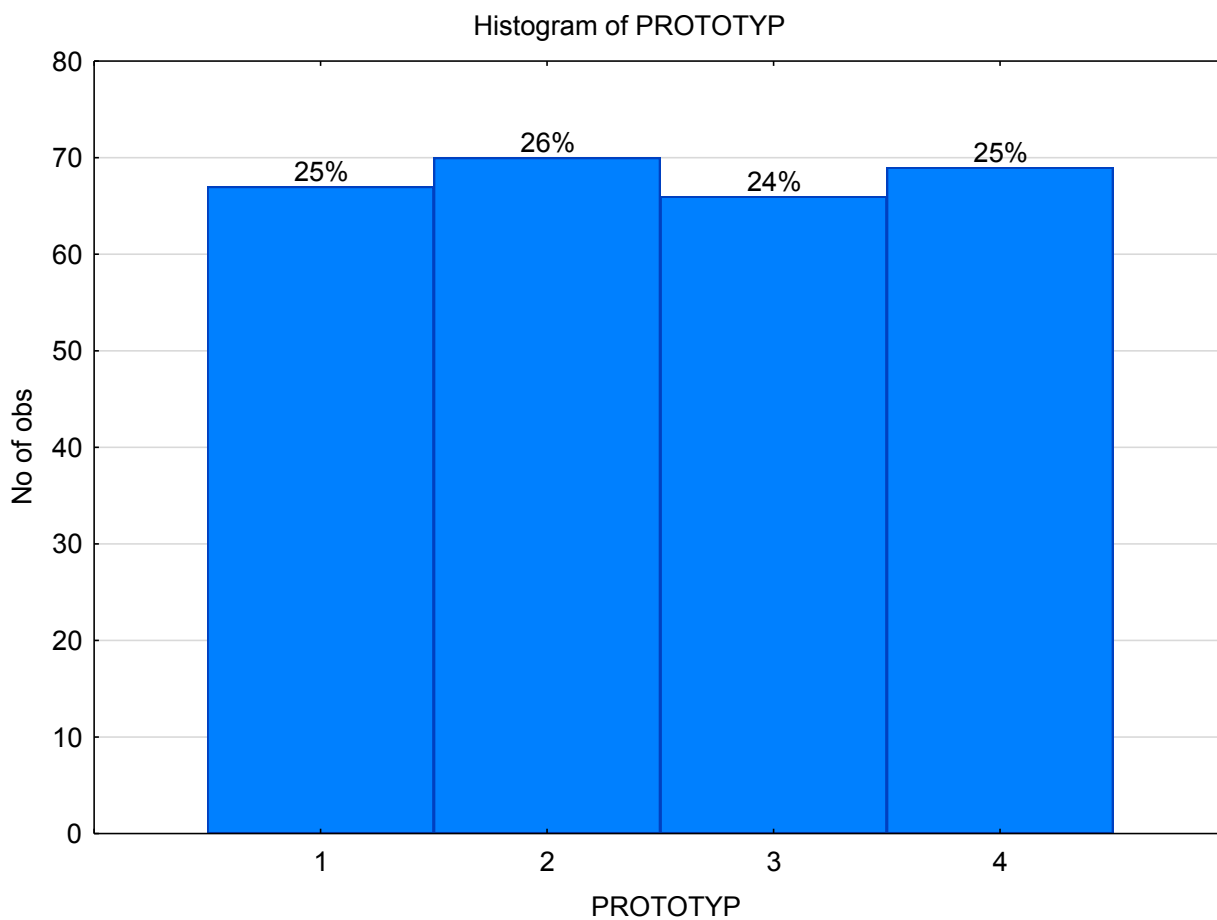
Referenz	Dokument	
Anhang 1	Datenauswertung: Deskriptive Statistik (auch online unter: <a href="https://bit.ly/MasterarbeitSBB">bit.ly/MasterarbeitSBB</a> )	
Anhang 2	Datenauswertung: Hypothesenvalidierung (auch online unter: <a href="https://bit.ly/MasterarbeitSBB">bit.ly/MasterarbeitSBB</a> )	
Anhang 3	Power-Berechnung	
Anhang 4	Fragebogen, letzte Version	
Anhang 5	Usability-Test: Szenario	
Anhang 6	Usability Tests: Pre-Interview	
Anhang 7	Usability Tests: Post-Interview	
Anhang 8	Usability Tests: Auswertung	
Anhang 9	Versuchsplan	
Anhang 10	Prototyp V1.0 und V2.1 (Screenshots)	



# Deskriptive Statistik

Für die Variable Prototyp

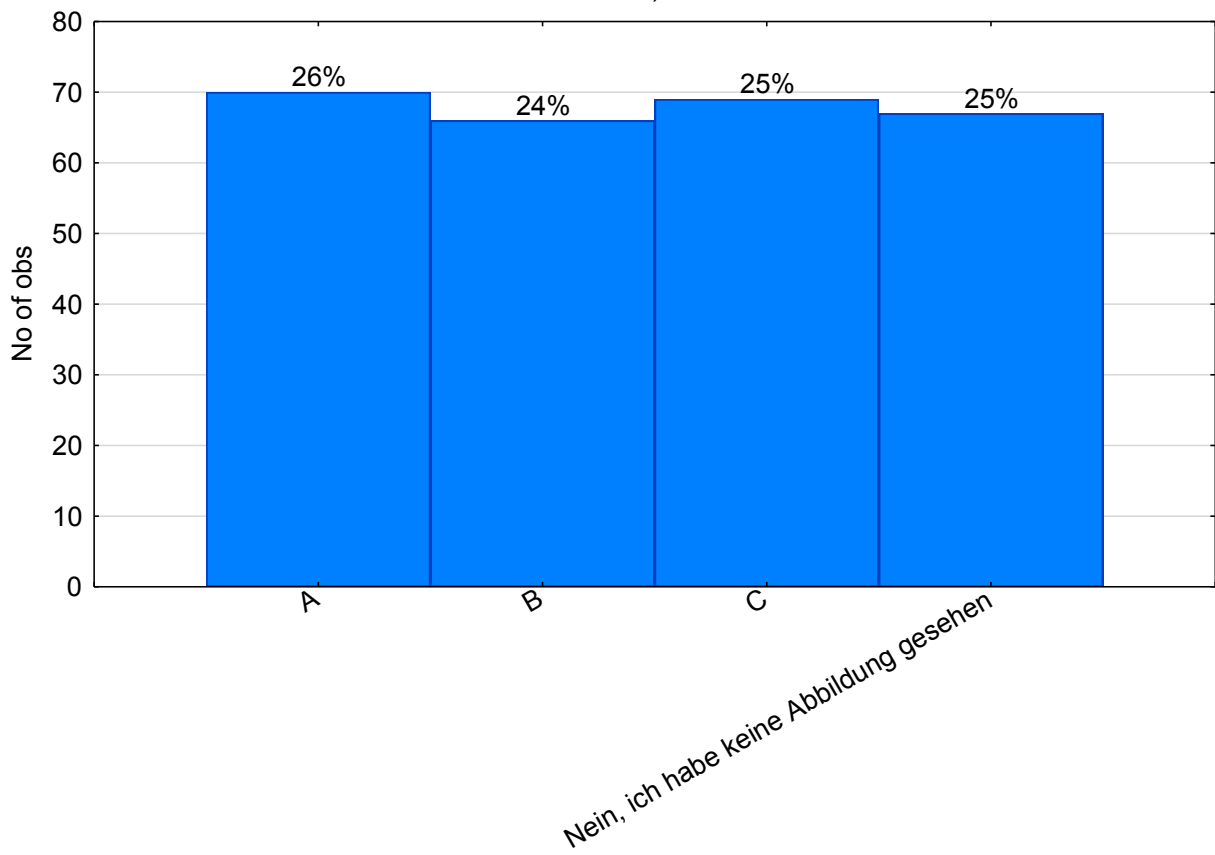
Category	Prototyp			
	Count	Cumulative Count	Percent	Cumulative Percent
1	67	67	24,63235	24,6324
2	70	137	25,73529	50,3676
3	66	203	24,26471	74,6324
4	69	272	25,36765	100,0000
Missing	0	272	0,00000	100,0000



Für die Variable „Welche Variable wahrgenommen?“

Category	Hast du eine Anzeige dieses Ladevorgangs gesehen? Falls ja, welche?			
	Count	Cumulative Count	Percent	Cumulative Percent
Nein, ich habe keine Abbildung gesehen	67	67	24,63235	24,6324
B	66	133	24,26471	48,8971
A	70	203	25,73529	74,6324
C	69	272	25,36765	100,0000
Missing	0	272	0,00000	100,0000

Histogram of 1. Hast du eine Anzeige dieses Ladevorgangs gesehen? Falls ja, welche? (siehe Abbild unten)



Für die Variable: Einschätzung der Ladezeit

Variable	Descriptive Statistics (Datensatz ohne leere Spalten - ohne Namen und E-Mail in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2. Wieviele Sekunden hat deiner Einschätzung n	272	2,410862E+09	5,000000	0,00	6,547546E+11	3,000000	6,000000	3,970015E+10

Einschätzung der Ladezeit, 3 Fälle (Ausreisser) ausgeschlossen

Va	All Groups Descriptive Statistics (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2.	269	5,799257	5,000000	0,00	45,00000	3,000000	6,000000	5,189383

nur Prototyp 1

Va	PROTOTYP=1 Descriptive Statistics (Spreadsheet in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2.	66	6,000000	5,000000	0,00	30,00000	3,000000	6,000000	5,400855

Nur Prototyp 2

Va	PROTOTYP=3 Descriptive Statistics (Spreadsheet in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2.	66	6,393939	5,000000	2,000000	45,00000	3,000000	7,000000	6,537181

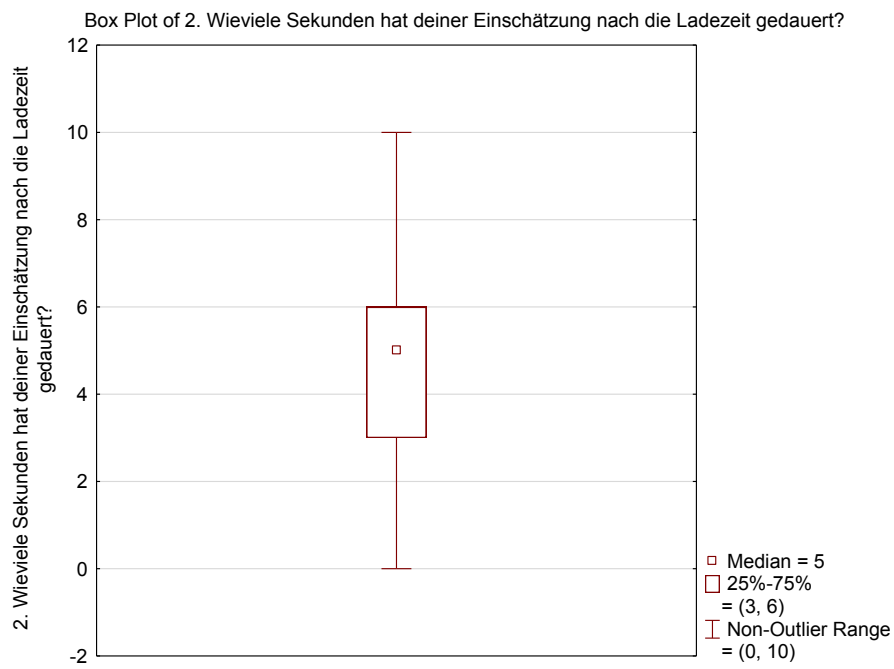
Nur Prototyp 3

Va	PROTOTYP=2 Descriptive Statistics (Spreadsheet in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2.	69	5,405797	5,000000	1,000000	40,00000	3,000000	5,000000	5,154435

Nur Prototyp 4

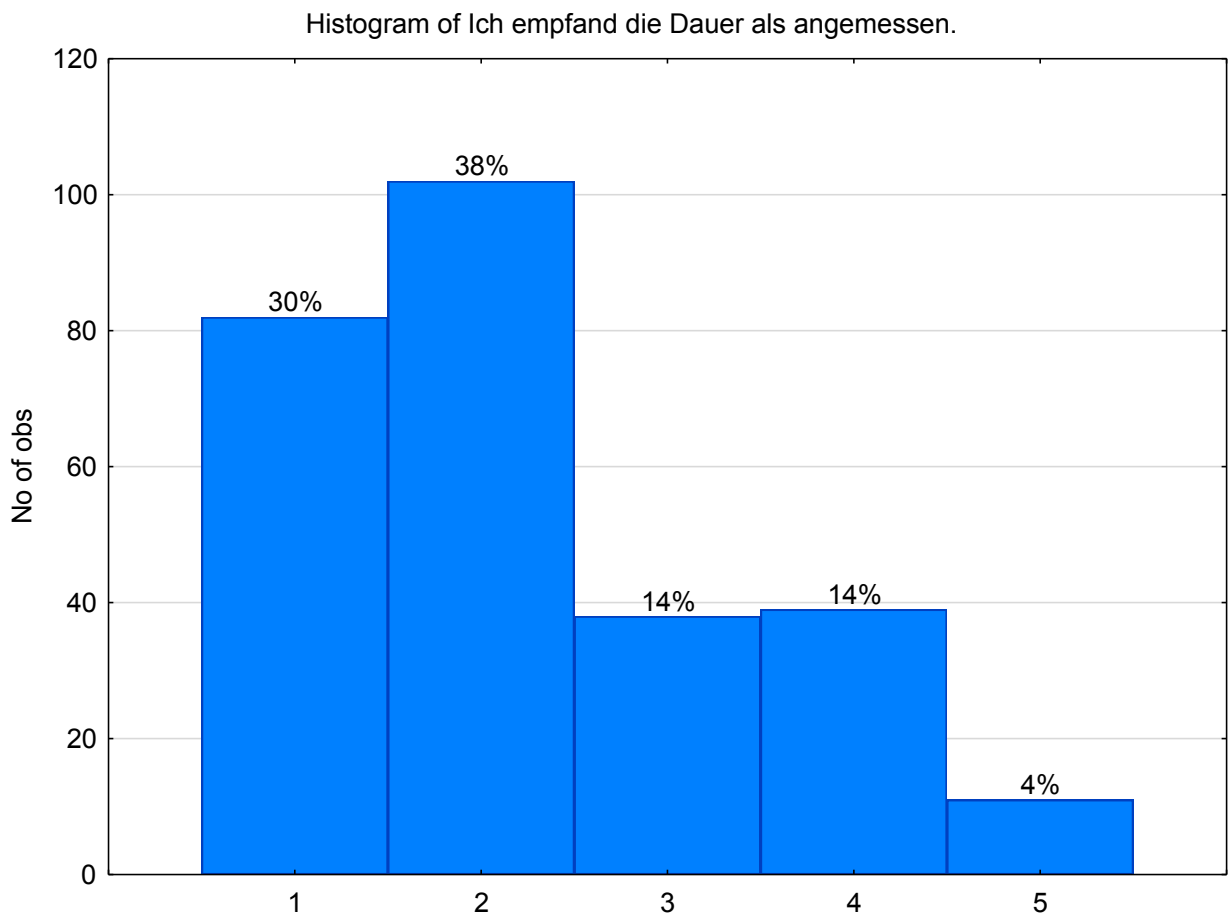
Va	PROTOTYP=4 Descriptive Statistics (Spreadsheet in Auswertung_SBB_Ladeanzeige)							
	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
2.	68	5,426471	5,000000	1,000000	20,00000	3,000000	6,000000	3,210599

## Ohne Ausreisser:



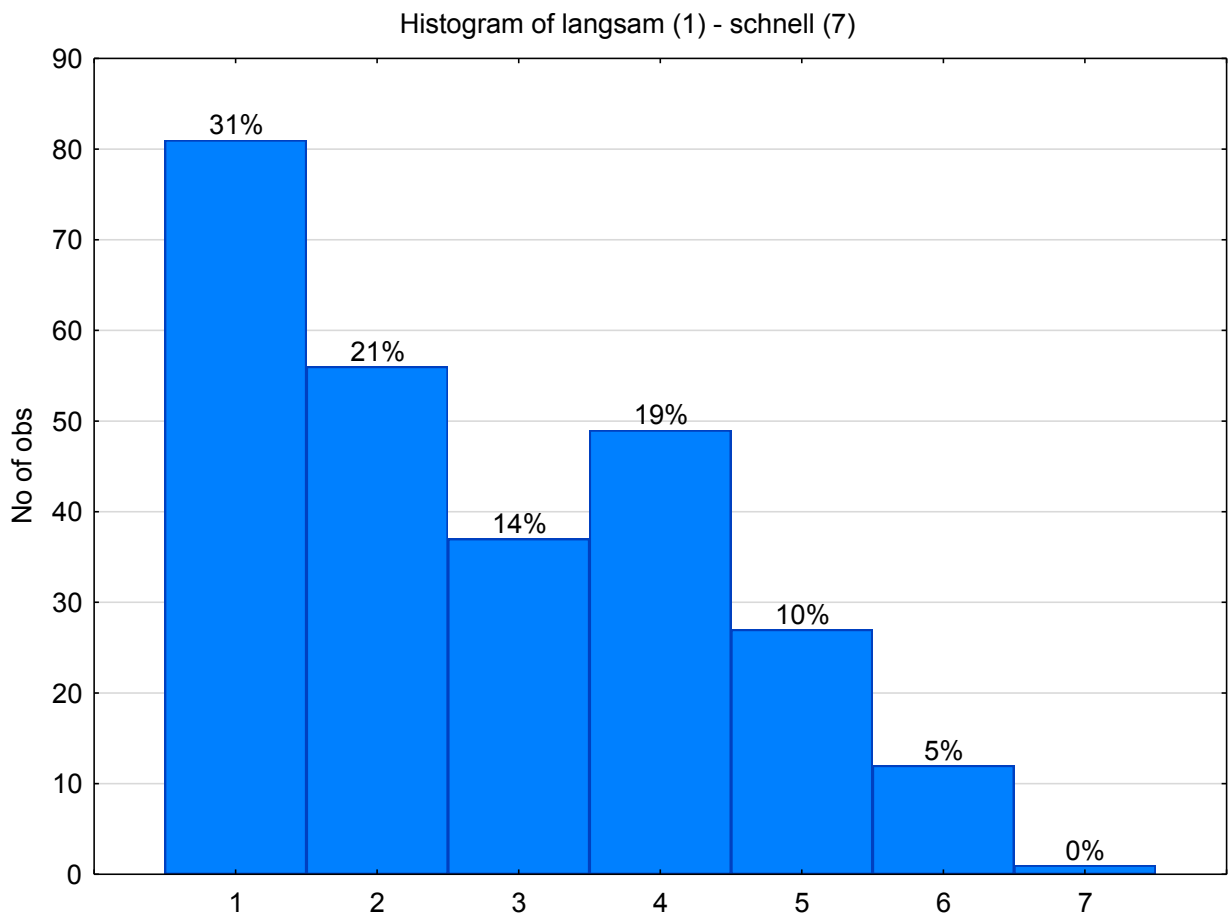
Für die Variable: empfundene Dauer

Category	Ich empfand die Dauer als angemessen			
	Count	Cumulative Count	Percent	Cumulative Percent
1	82	82	30,14706	30,1471
2	102	184	37,50000	67,6471
3	38	222	13,97059	81,6176
4	39	261	14,33824	95,9559
5	11	272	4,04412	100,0000
Missing	0	272	0,00000	100,0000



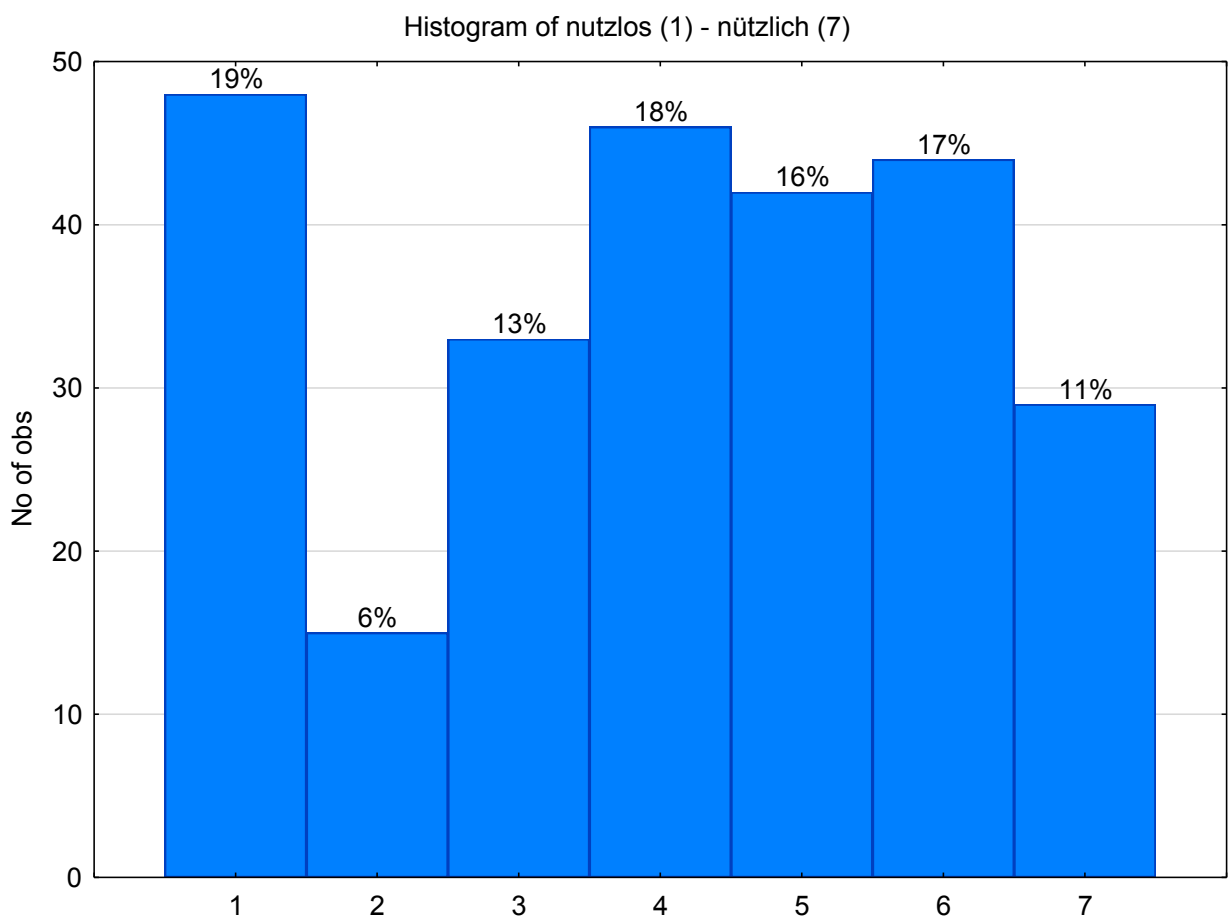
Für die Variable. Anzeige empfunden, schnell – langsam

Category	langsam (1) - schnell (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	81	81	29,77941	29,7794
2	56	137	20,58824	50,3676
3	37	174	13,60294	63,9706
4	49	223	18,01471	81,9853
5	27	250	9,92647	91,9118
6	12	262	4,41176	96,3235
7	1	263	0,36765	96,6912
Missing	9	272	3,30882	100,0000



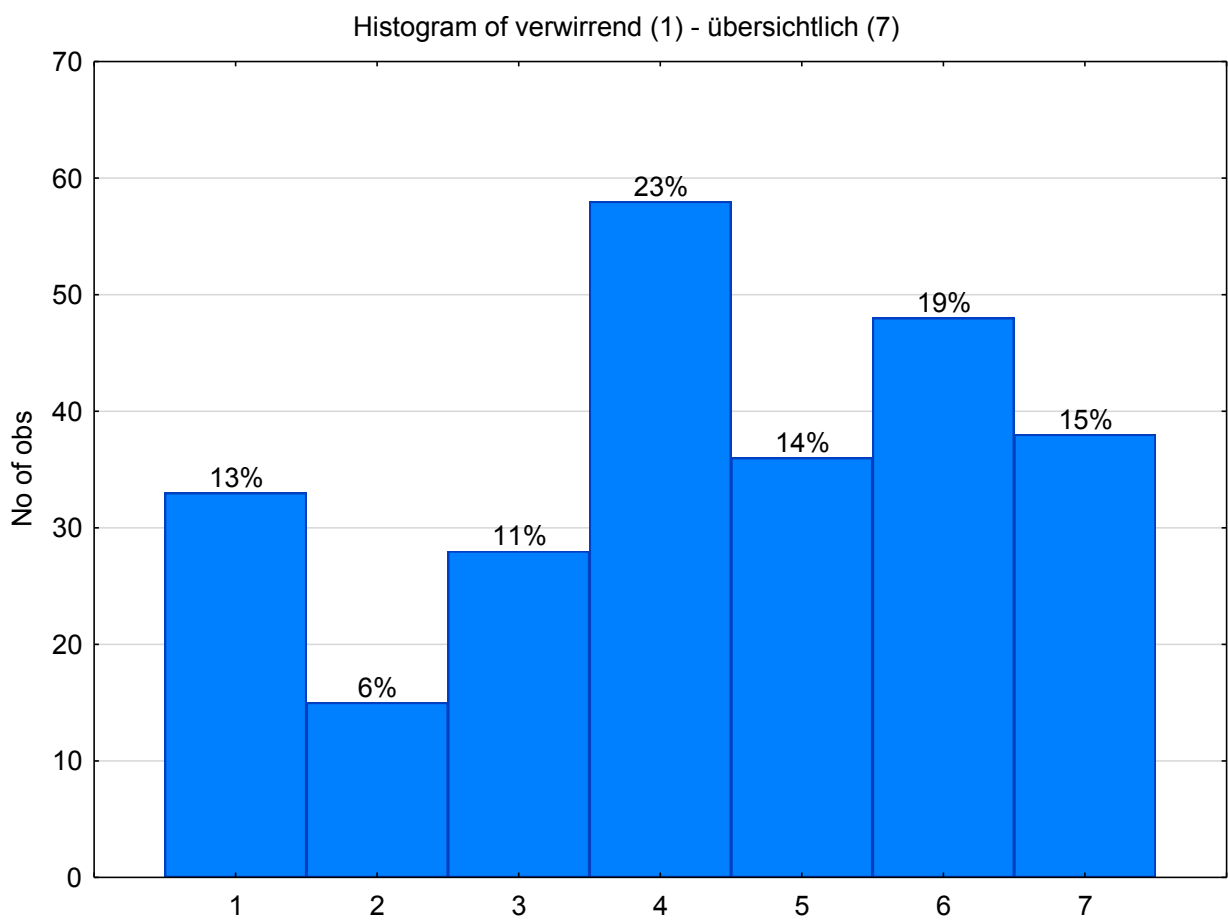
Für die Variable: Anzeige empfunden, nutzlos – nützlich

Category	nutzlos (1) - nützlich (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	48	48	17,64706	17,6471
2	15	63	5,51471	23,1618
3	33	96	12,13235	35,2941
4	46	142	16,91176	52,2059
5	42	184	15,44118	67,6471
6	44	228	16,17647	83,8235
7	29	257	10,66176	94,4853
Missing	15	272	5,51471	100,0000



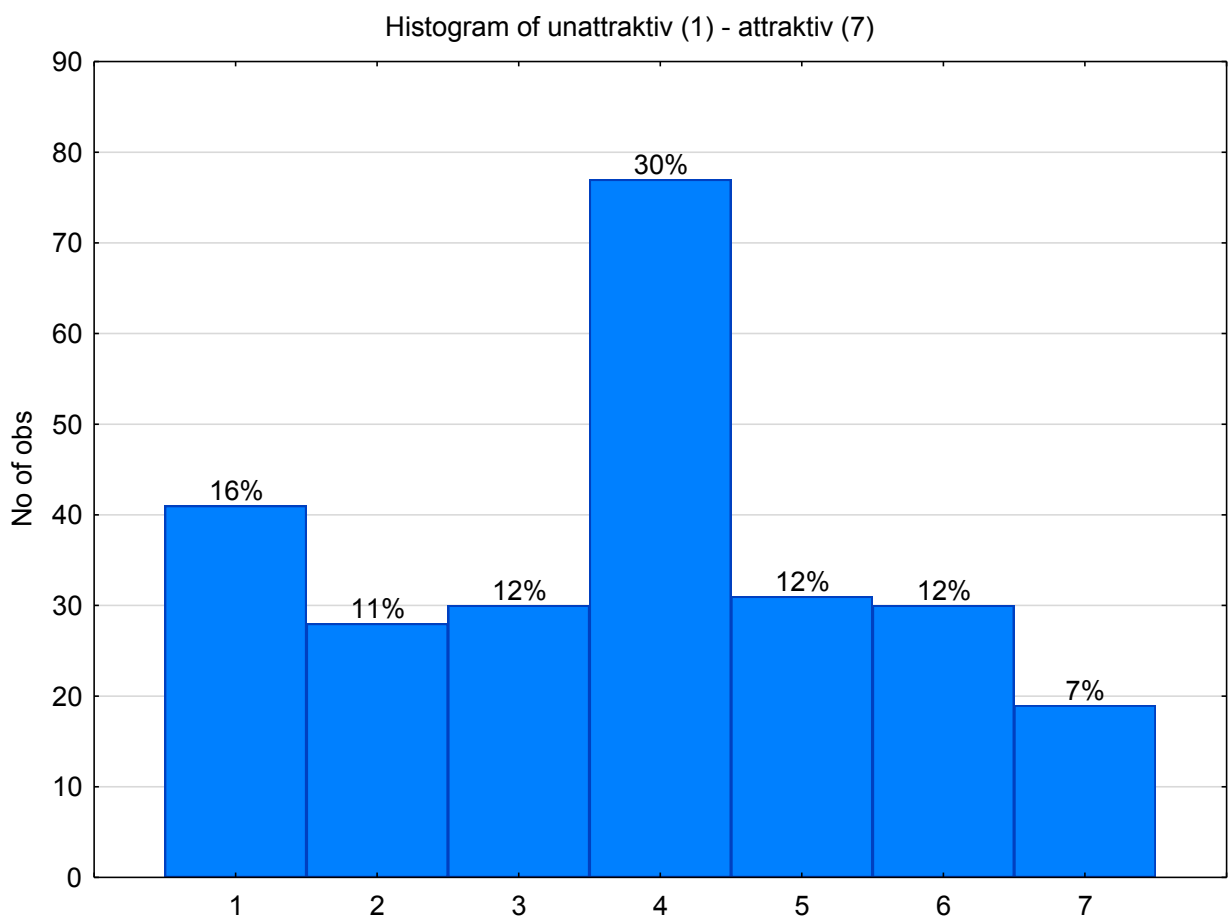
Für die Variable: Anzeige empfunden, verwirrend – übersichtlich

Category	verwirrend (1) - übersichtlich (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	33	33	12,13235	12,1324
2	15	48	5,51471	17,6471
3	28	76	10,29412	27,9412
4	58	134	21,32353	49,2647
5	36	170	13,23529	62,5000
6	48	218	17,64706	80,1471
7	38	256	13,97059	94,1176
Missing	16	272	5,88235	100,0000



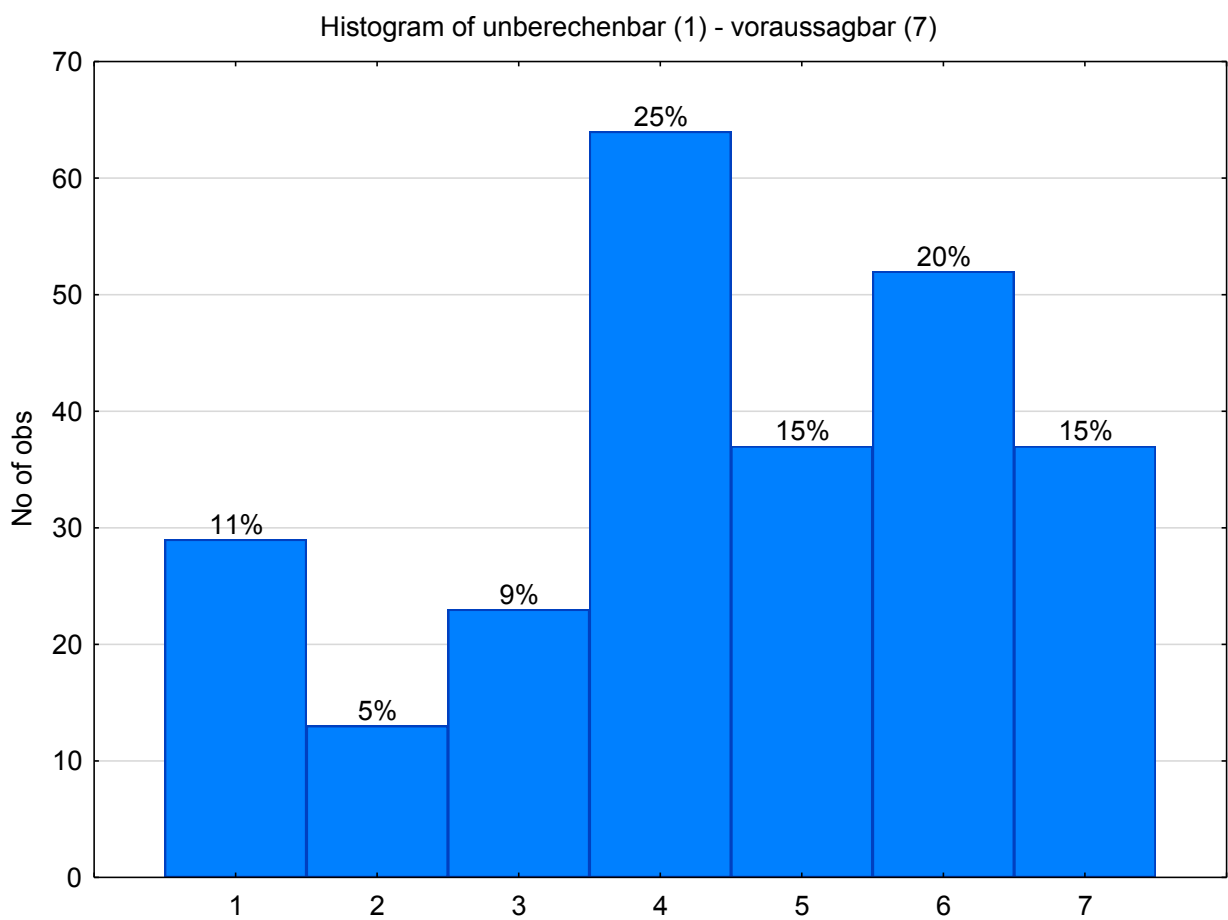
Für die Variable: Anzeige empfunden, unattraktiv – attraktiv

Category	unattraktiv (1) - attraktiv (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	41	41	15,07353	15,0735
2	28	69	10,29412	25,3676
3	30	99	11,02941	36,3971
4	77	176	28,30882	64,7059
5	31	207	11,39706	76,1029
6	30	237	11,02941	87,1324
7	19	256	6,98529	94,1176
Missing	16	272	5,88235	100,0000



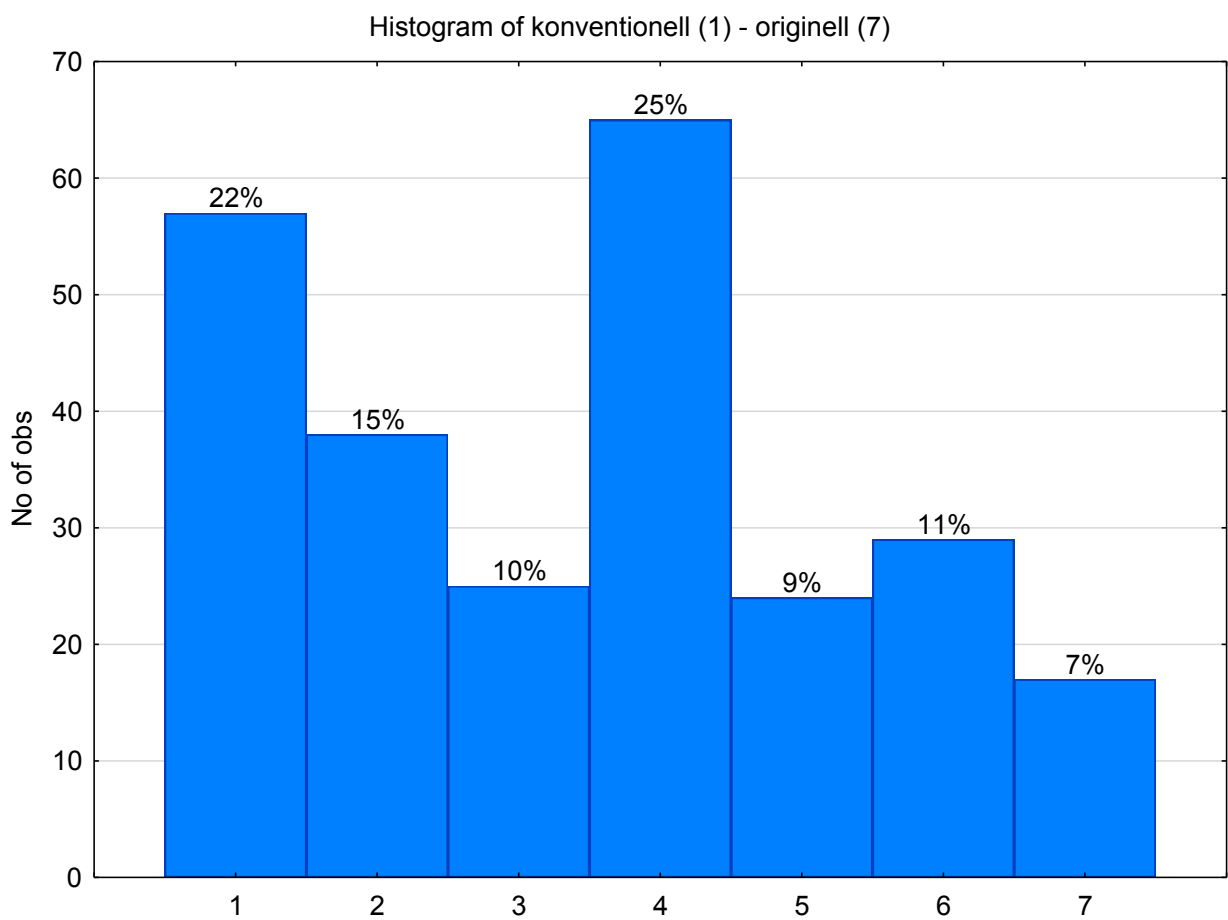
Für die Variable: Anzeige empfunden, unberechenbar – voraussagbar

Category	unberechenbar (1) - voraussagbar (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	29	29	10,66176	10,6618
2	13	42	4,77941	15,4412
3	23	65	8,45588	23,8971
4	64	129	23,52941	47,4265
5	37	166	13,60294	61,0294
6	52	218	19,11765	80,1471
7	37	255	13,60294	93,7500
Missing	17	272	6,25000	100,0000



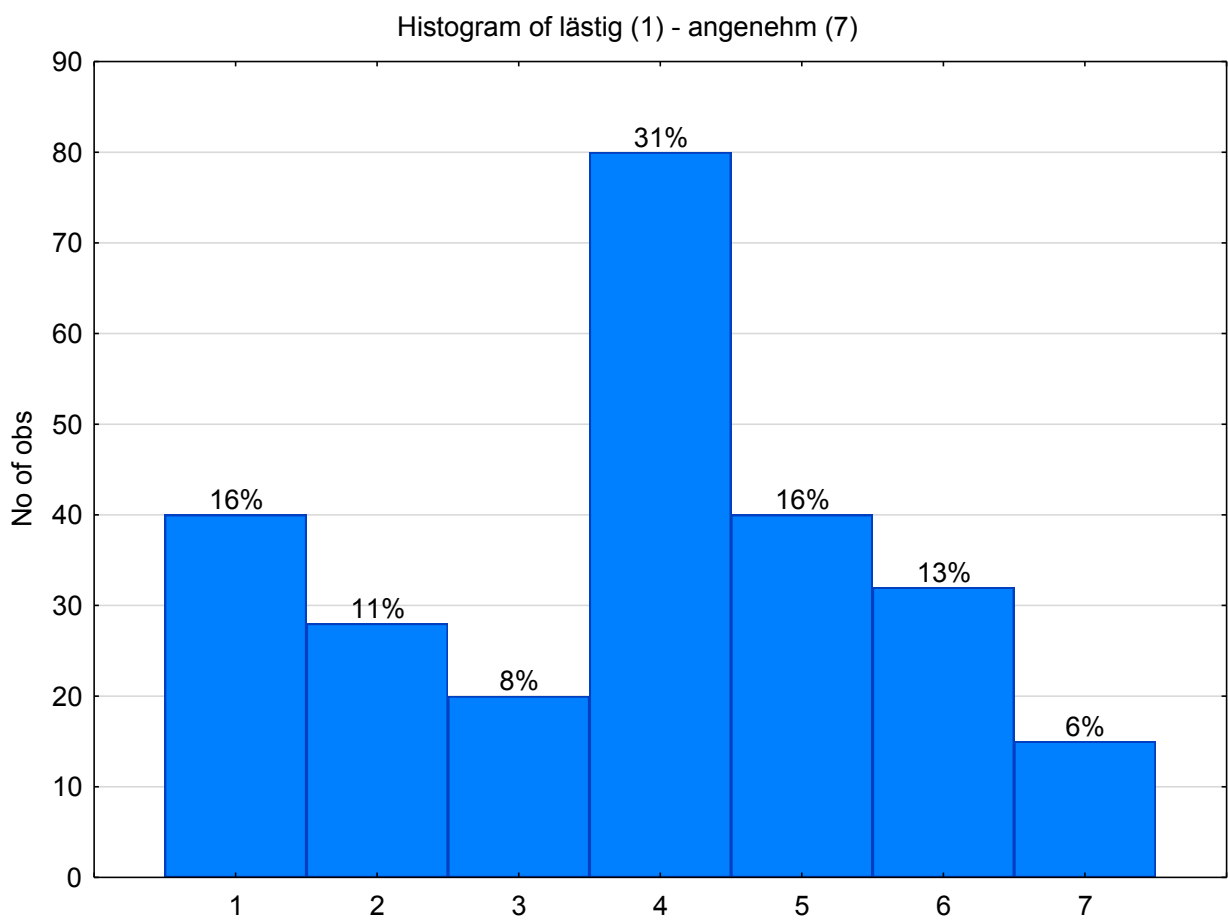
Für die Variable: Anzeige empfunden, konventionell – originell

Category	konventionell (1) - originell (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	57	57	20,95588	20,9559
2	38	95	13,97059	34,9265
3	25	120	9,19118	44,1176
4	65	185	23,89706	68,0147
5	24	209	8,82353	76,8382
6	29	238	10,66176	87,5000
7	17	255	6,25000	93,7500
Missing	17	272	6,25000	100,0000



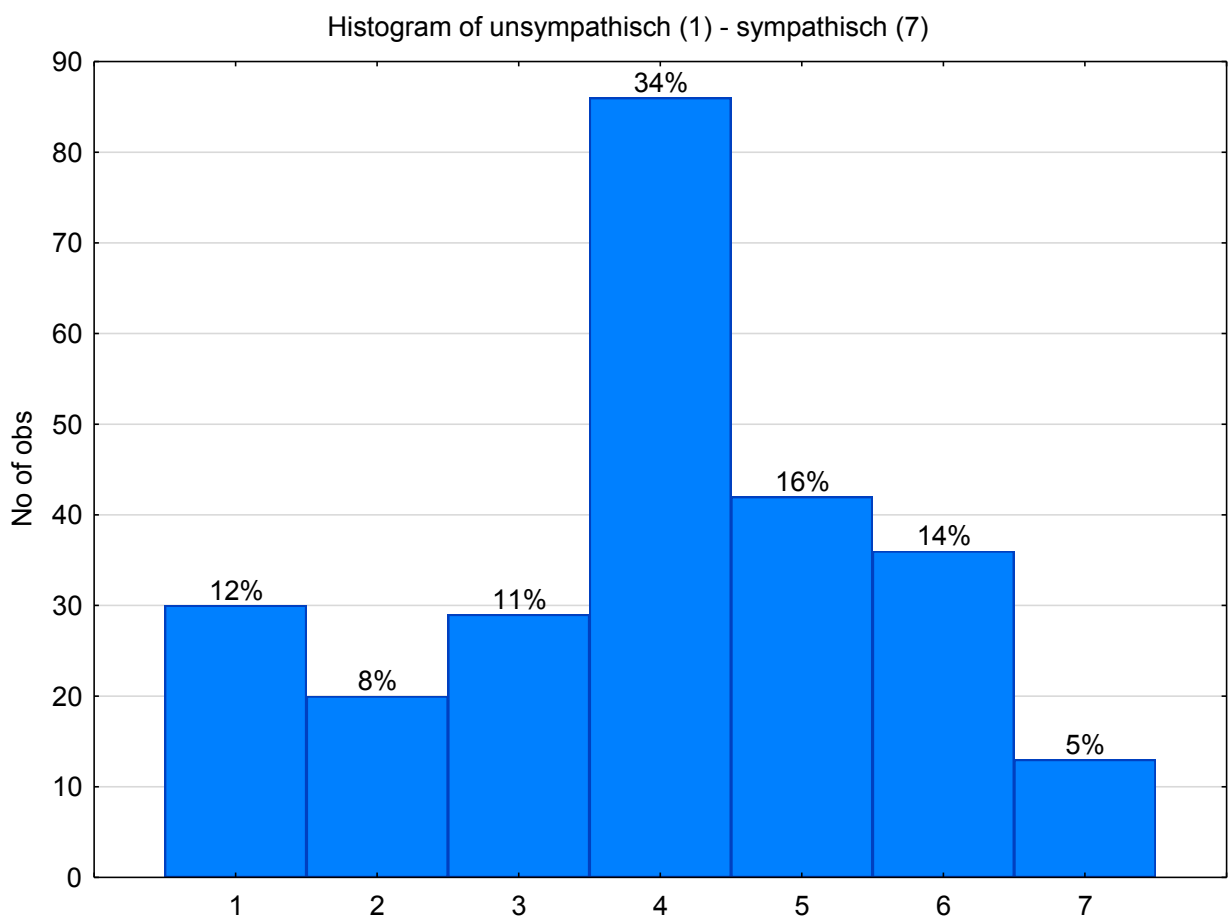
Für die Variable: Anzeige empfunden, lästig – angenehm

Category	lästig (1) - angenehm (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	40	40	14,70588	14,7059
2	28	68	10,29412	25,0000
3	20	88	7,35294	32,3529
4	80	168	29,41176	61,7647
5	40	208	14,70588	76,4706
6	32	240	11,76471	88,2353
7	15	255	5,51471	93,7500
Missing	17	272	6,25000	100,0000



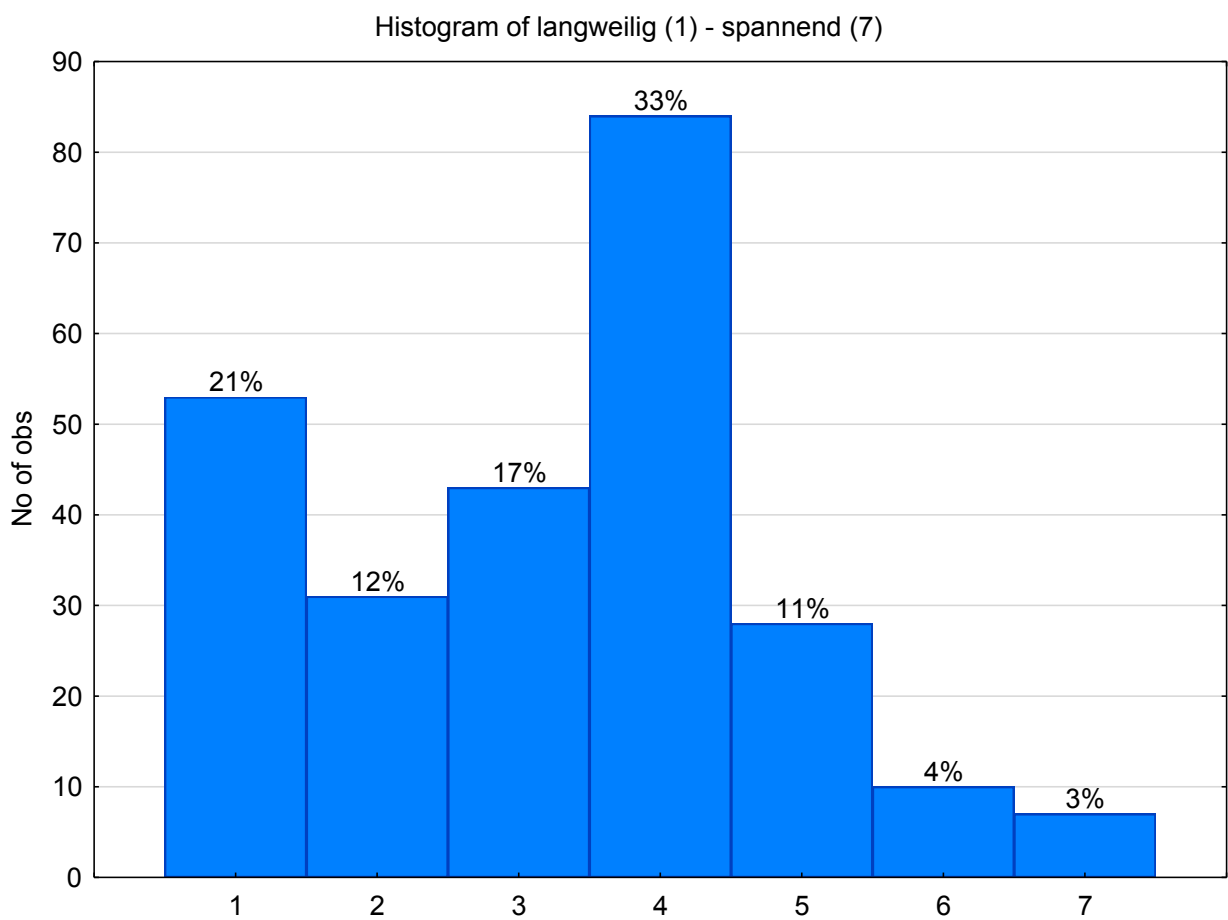
Für die Variable: Anzeige empfunden, unsympathisch – sympathisch

Category	unsympathisch (1) - sympathisch (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	30	30	11,02941	11,02941
2	20	50	7,35294	18,3824
3	29	79	10,66176	29,0441
4	86	165	31,61765	60,6618
5	42	207	15,44118	76,1029
6	36	243	13,23529	89,3382
7	13	256	4,77941	94,1176
Missing	16	272	5,88235	100,0000



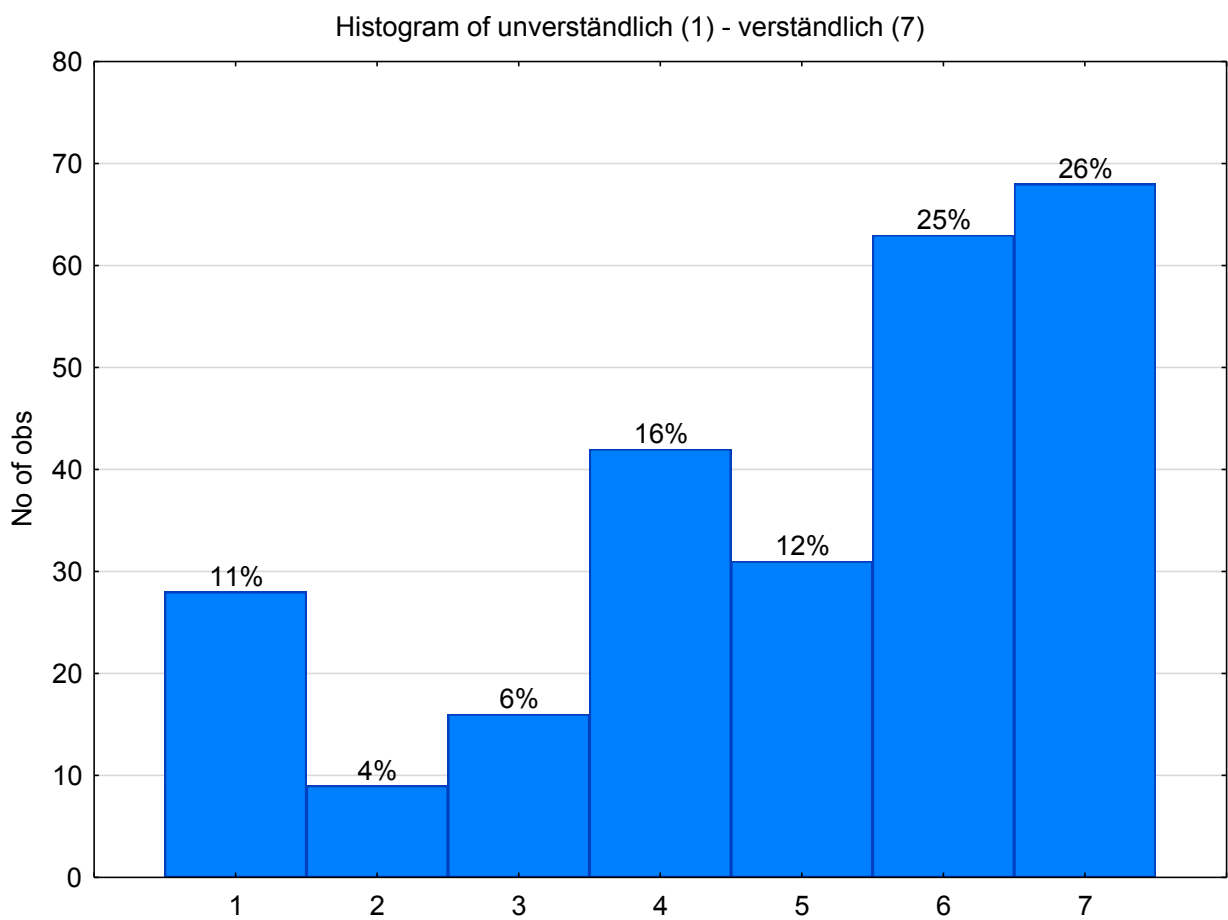
Für die Variable: Anzeige empfunden, langweilig – spannend

Category	langweilig (1) - spannend (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	53	53	19,48529	19,4853
2	31	84	11,39706	30,8824
3	43	127	15,80882	46,6912
4	84	211	30,88235	77,5735
5	28	239	10,29412	87,8676
6	10	249	3,67647	91,5441
7	7	256	2,57353	94,1176
Missing	16	272	5,88235	100,0000



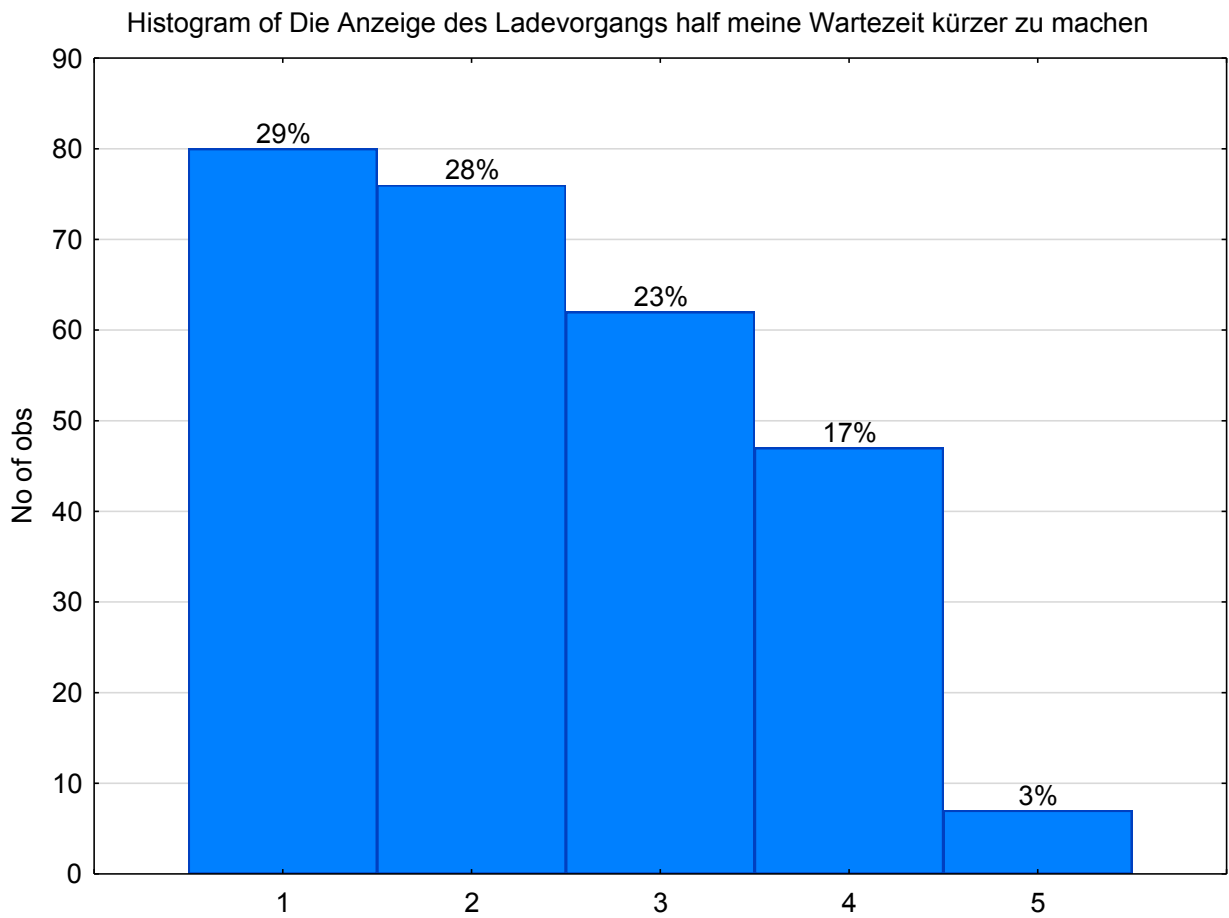
Für die Variable: Anzeige empfunden, unverständlich – verständlich

Category	unverständlich (1) - verständlich (7)			
	Count	Cumulative Count	Percent	Cumulative Percent
1	28	28	10,29412	10,2941
2	9	37	3,30882	13,6029
3	16	53	5,88235	19,4853
4	42	95	15,44118	34,9265
5	31	126	11,39706	46,3235
6	63	189	23,16176	69,4853
7	68	257	25,00000	94,4853
Missing	15	272	5,51471	100,0000



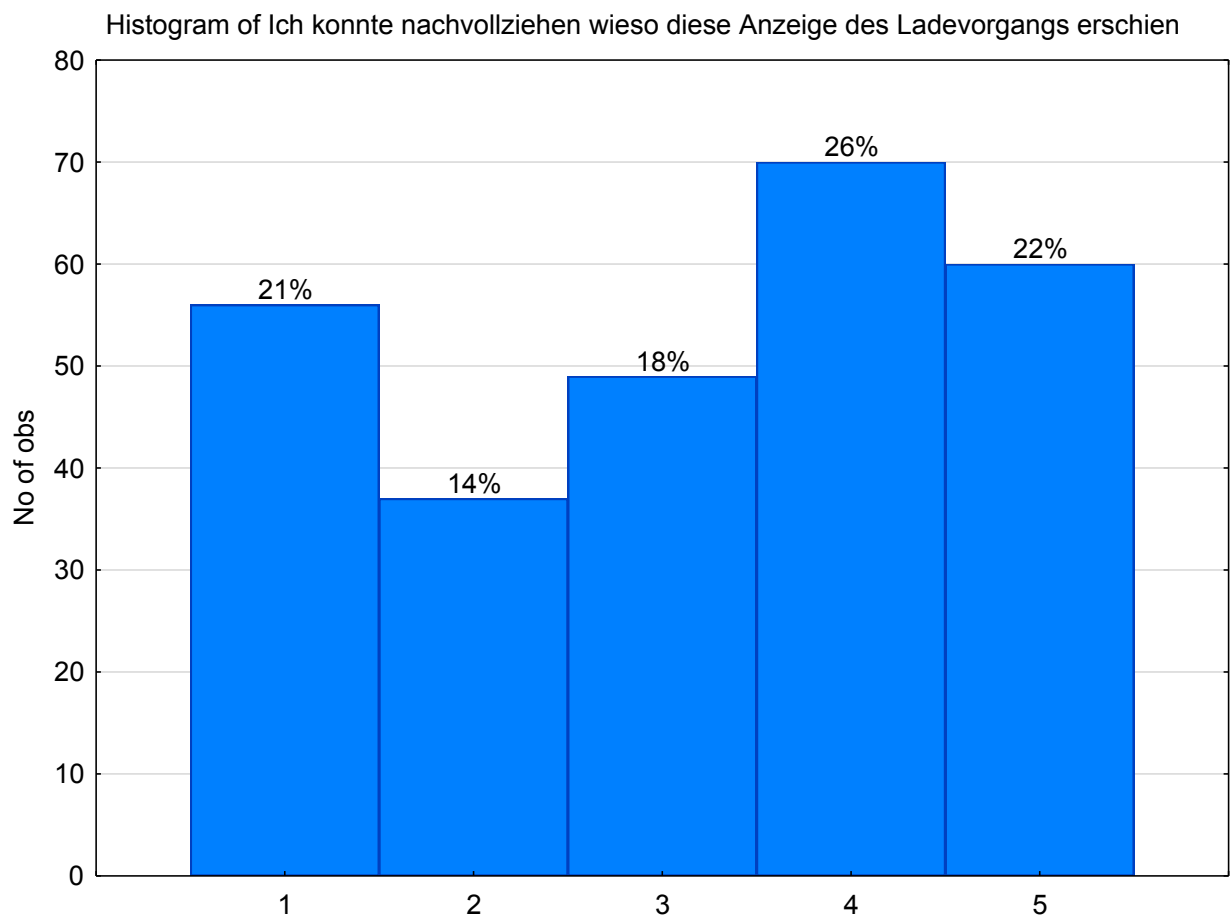
Für die Variable: „Die Anzeige half die Wartezeit kürzer zu machen“

Category	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen			
	Count	Cumulative Count	Percent	Cumulative Percent
1	80	80	29,41176	29,4118
2	76	156	27,94118	57,3529
3	62	218	22,79412	80,1471
4	47	265	17,27941	97,4265
5	7	272	2,57353	100,0000
Missing	0	272	0,00000	100,0000



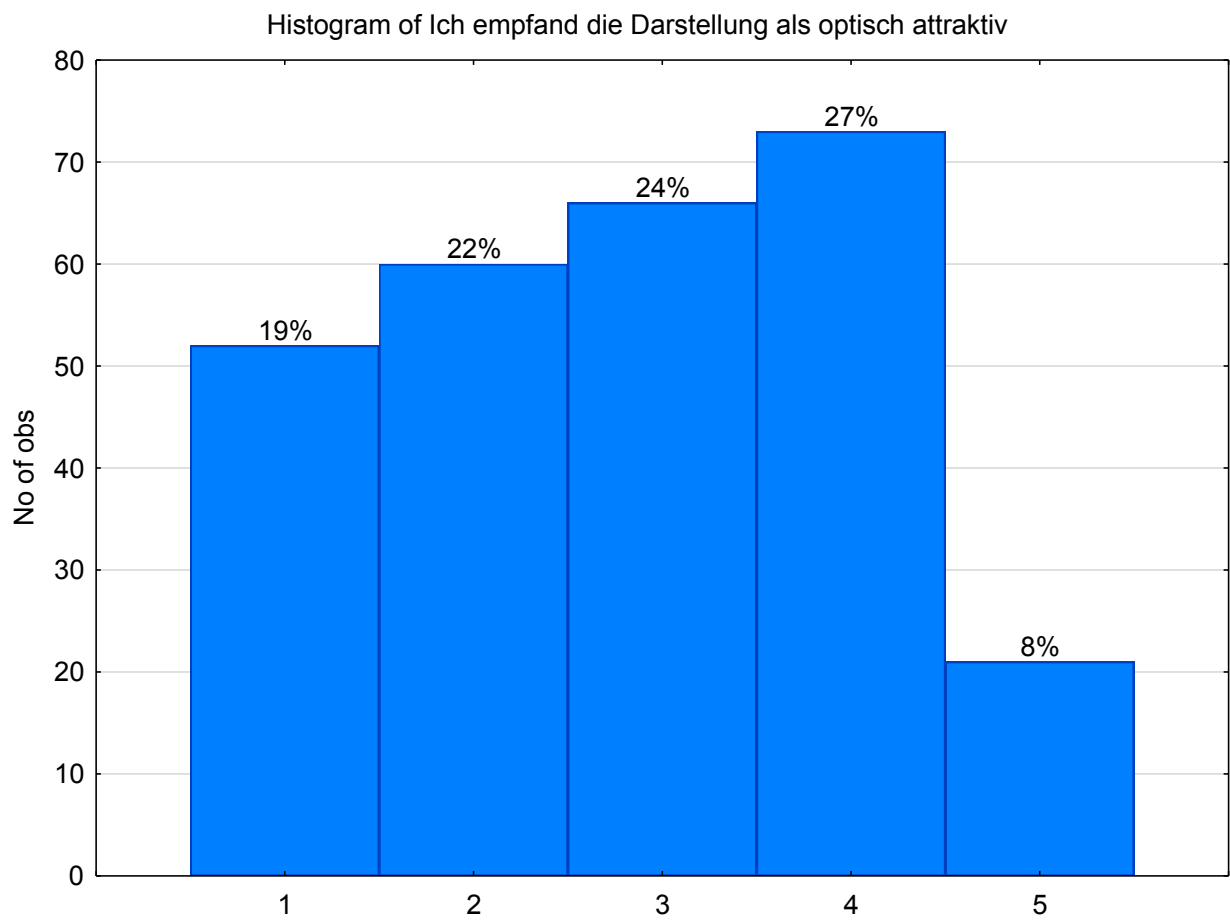
Für die Variable: „Anzeige des Ladevorgangs nachvollziehbar“

Category	Ich konnte nachvollziehen wieso diese Anzeige des Ladevorgangs			
	Count	Cumulative Count	Percent	Cumulative Percent
1	56	56	20,58824	20,5882
2	37	93	13,60294	34,1912
3	49	142	18,01471	52,2059
4	70	212	25,73529	77,9412
5	60	272	22,05882	100,0000
Missing	0	272	0,00000	100,0000



Für die Variable „Anzeige optisch attraktiv“

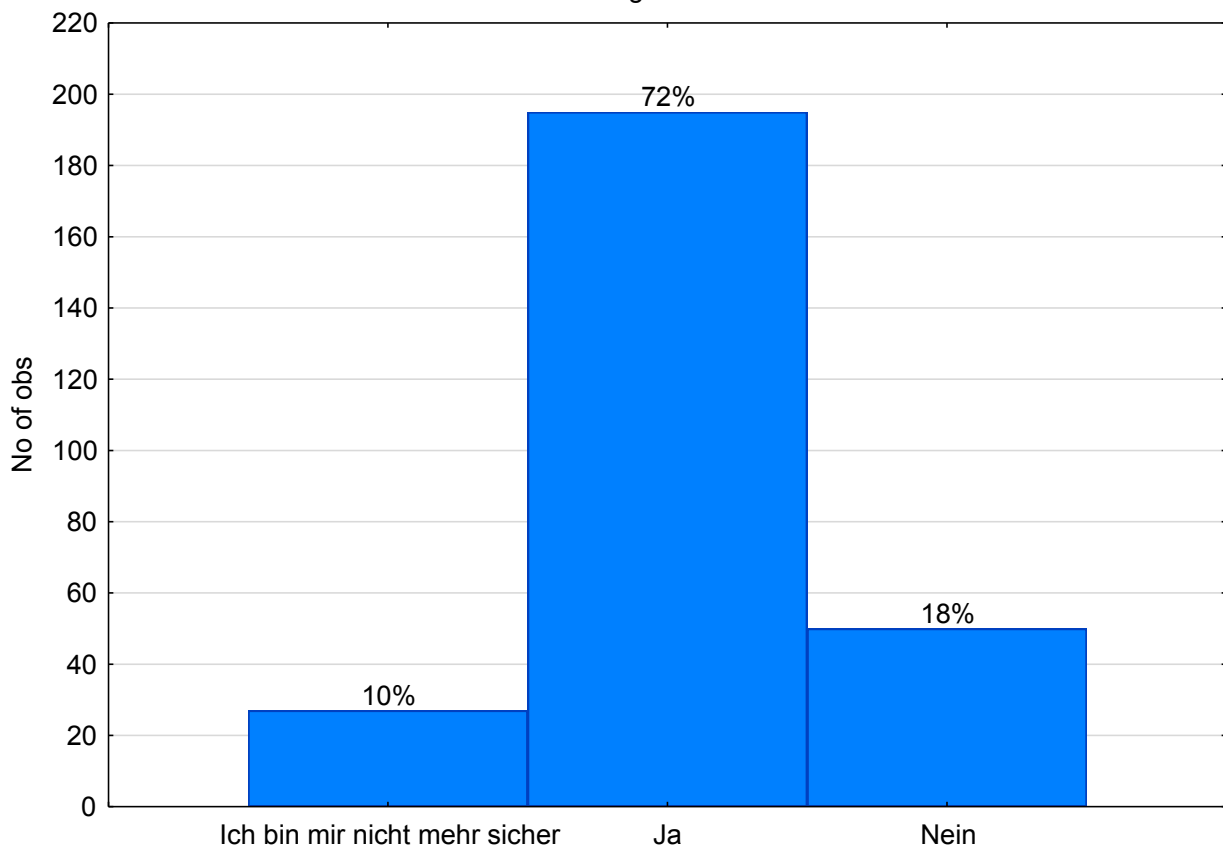
Category	Ich empfand die Darstellung als optisch attraktiv			
	Count	Cumulative Count	Percent	Cumulative Percent
1	52	52	19,11765	19,11765
2	60	112	22,05882	41,17647
3	66	178	24,26471	65,44118
4	73	251	26,83824	92,27941
5	21	272	7,72059	100,00000
Missing	0	272	0,00000	100,00000



Für die Variable „Hat sich der Warenkorb aktualisiert?“

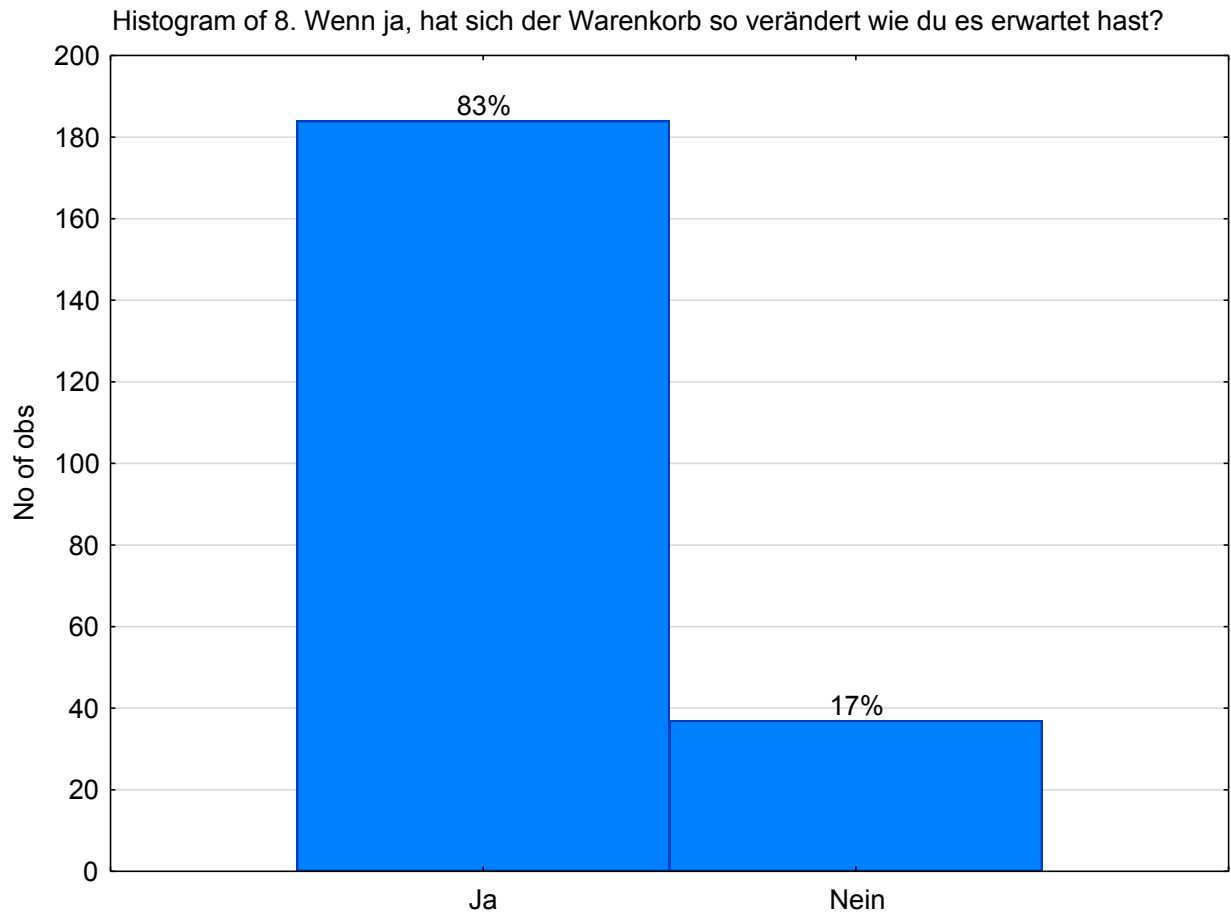
Category	Während des Ladevorgangs hat sich der Warenkorb aktualisiert			
	Count	Cumulative Count	Percent	Cumulative Percent
Ja	195	195	71,69118	71,6912
Ich bin mir nicht mehr sicher	27	222	9,92647	81,6176
Nein	50	272	18,38235	100,0000
Missing	0	272	0,00000	100,0000

Histogram of 7. Während des Ladevorgangs hat sich der Warenkorb aktualisiert. Hast du diese Veränderung bemerkt?



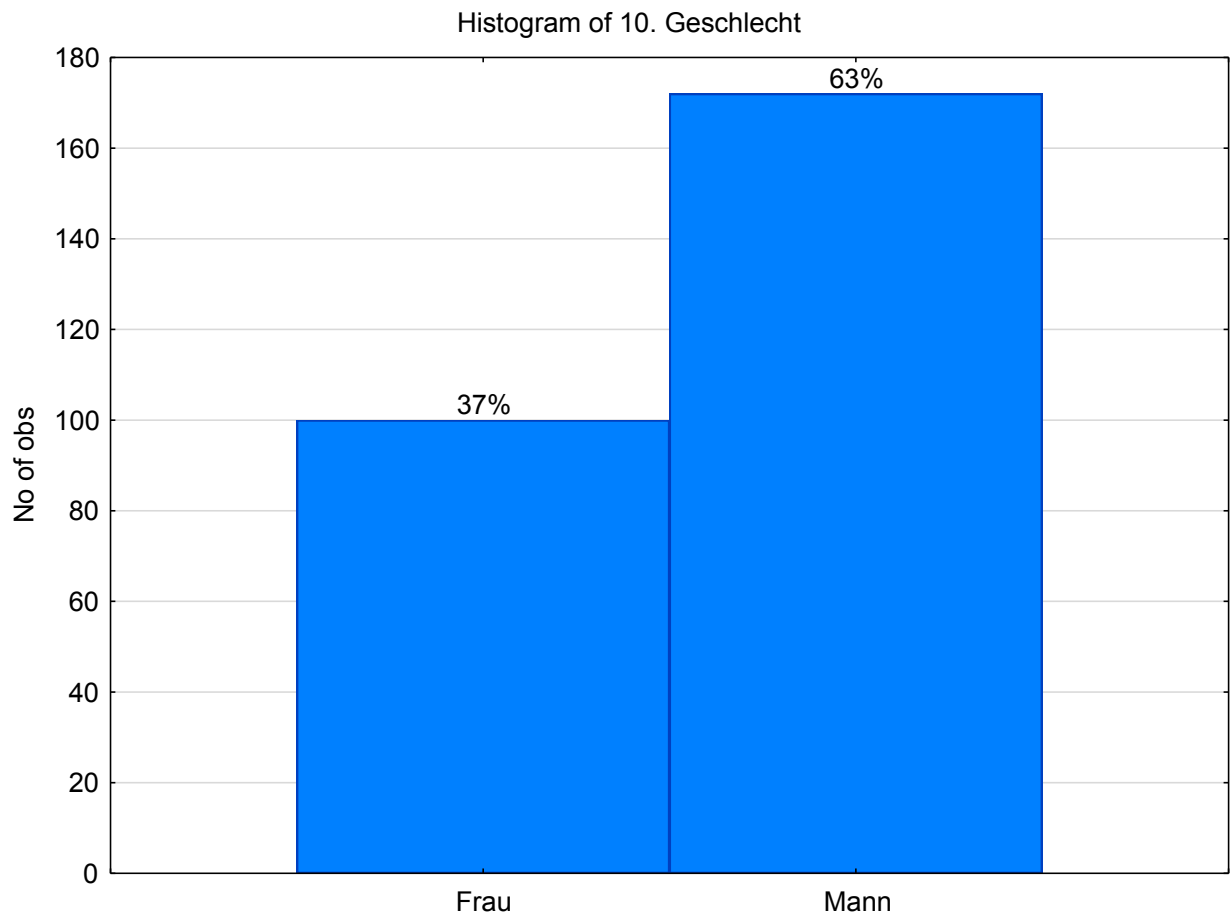
Für die Variable: „Veränderung wie erwartet?“

Category	Wenn ja, hat sich der Warenkorb so verändert wie			
	Count	Cumulative Count	Percent	Cumulative Percent
Ja	184	184	67,64706	67,647
Nein	37	221	13,60294	81,250
Missing	51	272	18,75000	100,000



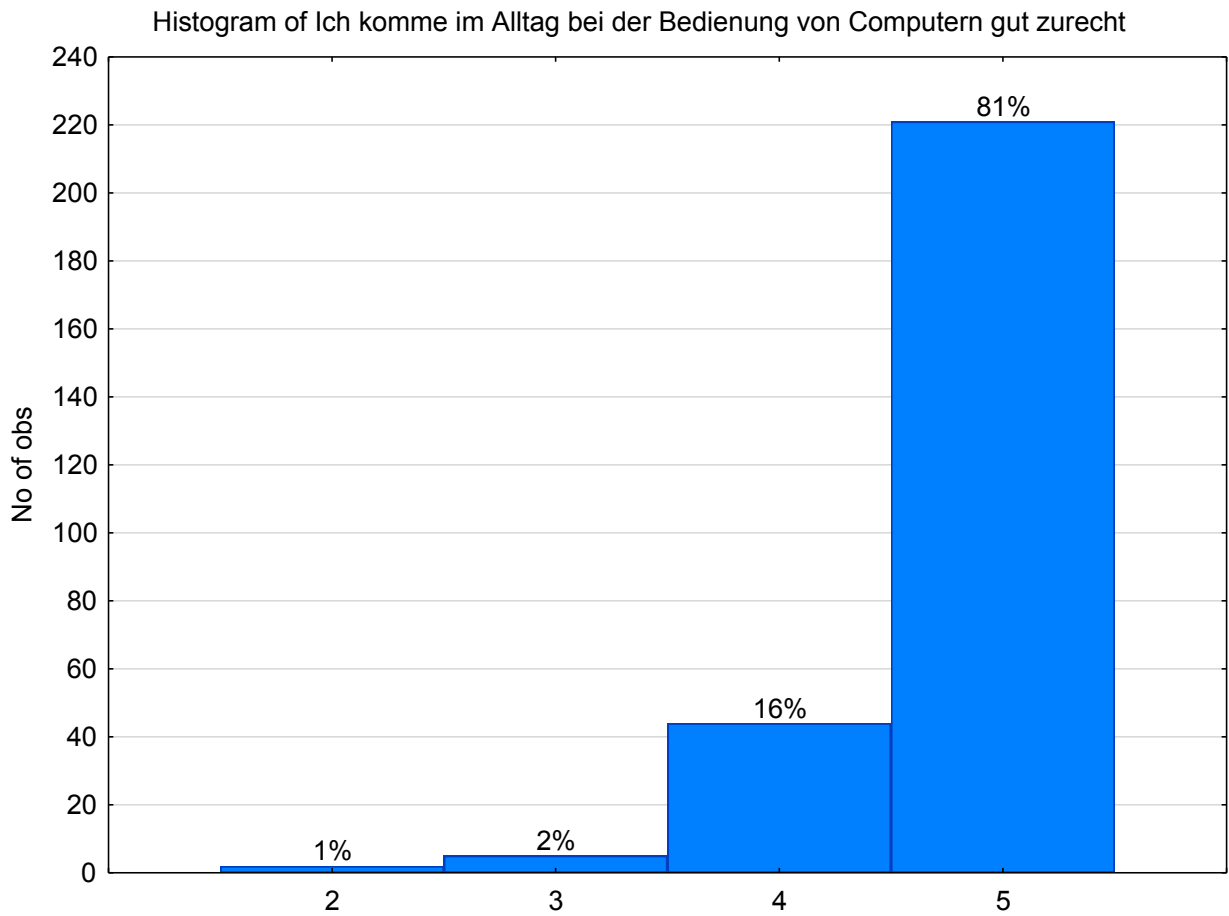
Für die Variable: Geschlecht

Category	Geschlecht			
	Count	Cumulative Count	Percent	Cumulative Percent
Mann	172	172	63,23529	63,23529
Frau	100	272	36,76471	100,00000
Missing	0	272	0,00000	100,00000



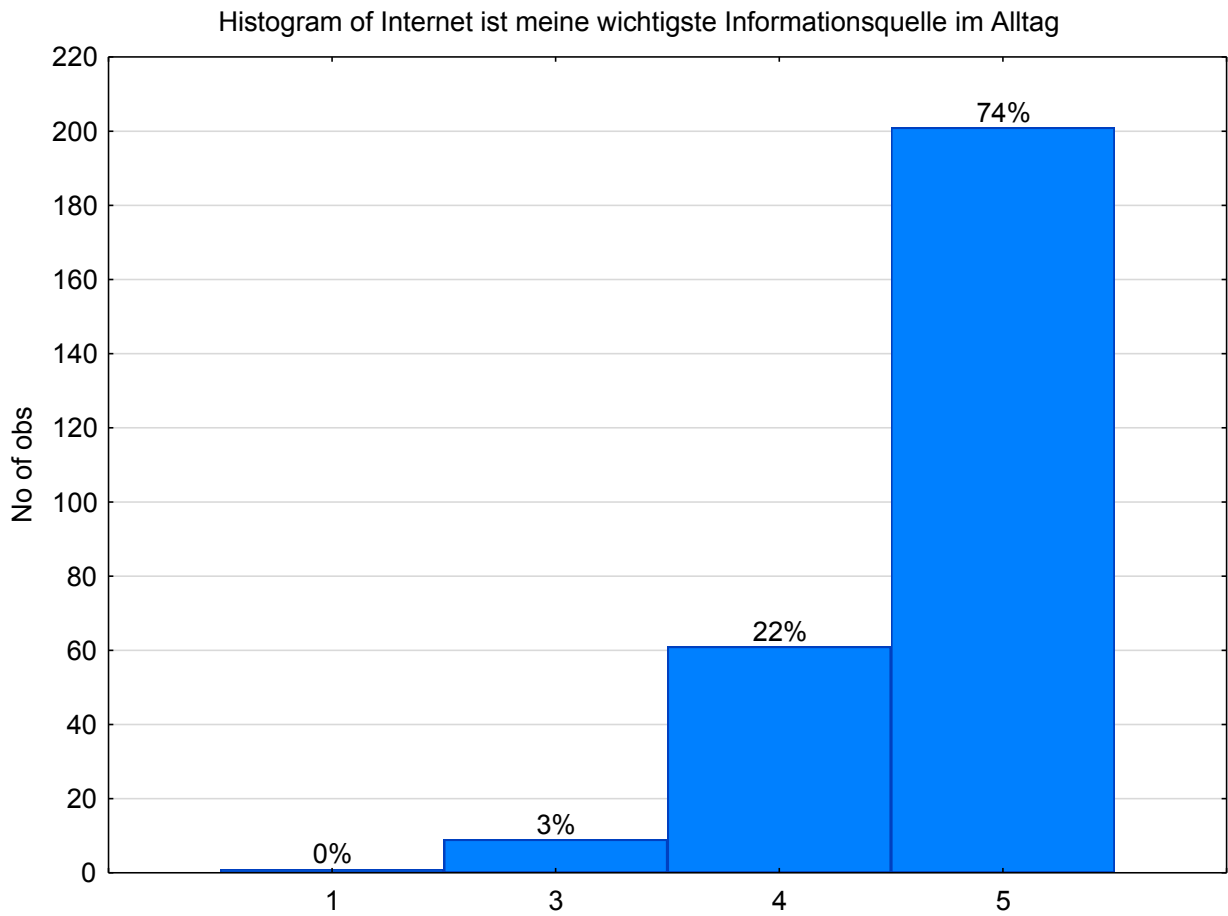
Für die Variable: „Zurechtkommen mit Computern im Alltag?“

Category	Ich komme im Alltag bei der Bedienung von Computern gut zurecht			
	Count	Cumulative Count	Percent	Cumulative Percent
2	2	2	0,73529	0,7353
3	5	7	1,83824	2,5735
4	44	51	16,17647	18,7500
5	221	272	81,25000	100,0000
Missing	0	272	0,00000	100,0000



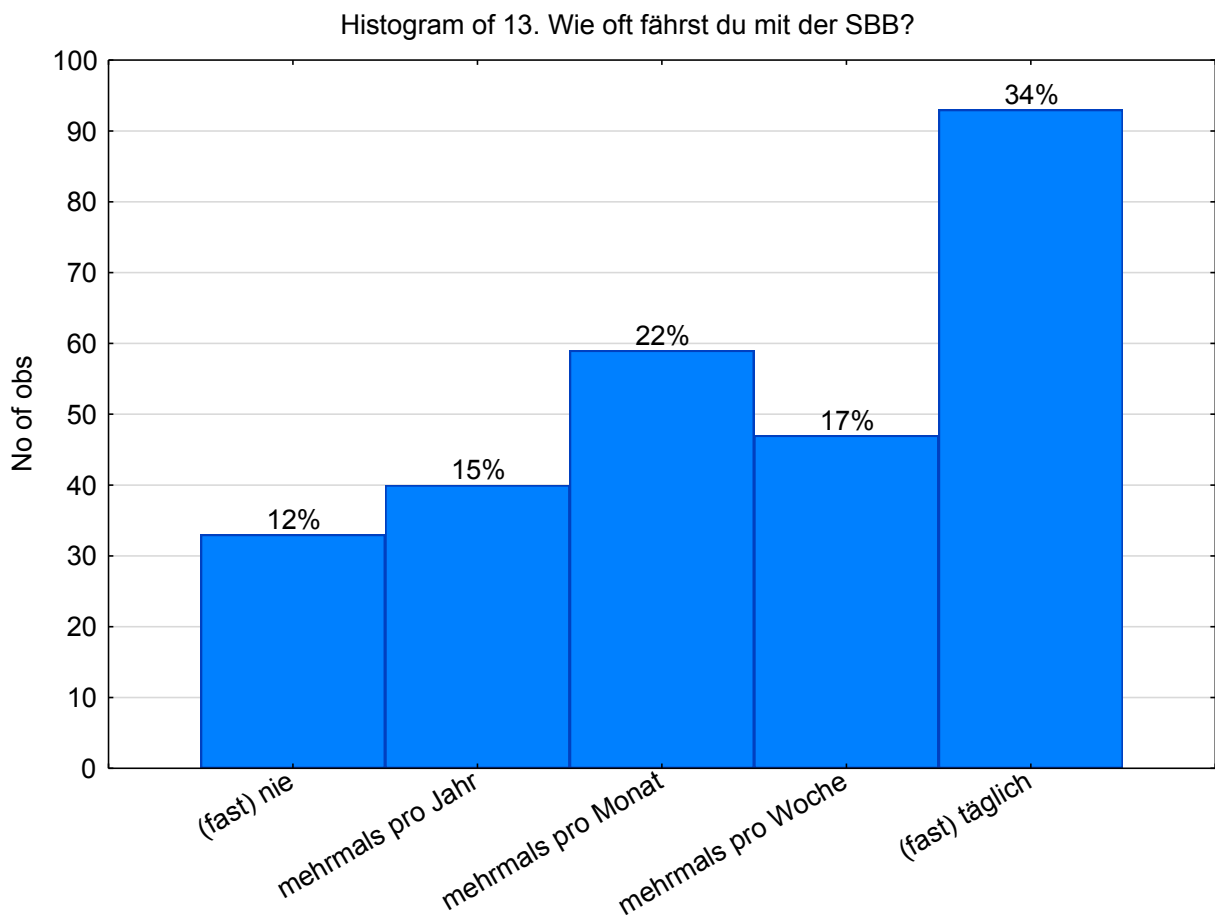
Für die Variable: „Internet als wichtige Informationsquelle“

Category	Internet ist meine wichtigste Informationsquelle im Alltag			
	Count	Cumulative Count	Percent	Cumulative Percent
1	1	1	0,36765	0,3676
3	9	10	3,30882	3,6765
4	61	71	22,42647	26,1029
5	201	272	73,89706	100,0000
Missing	0	272	0,00000	100,0000



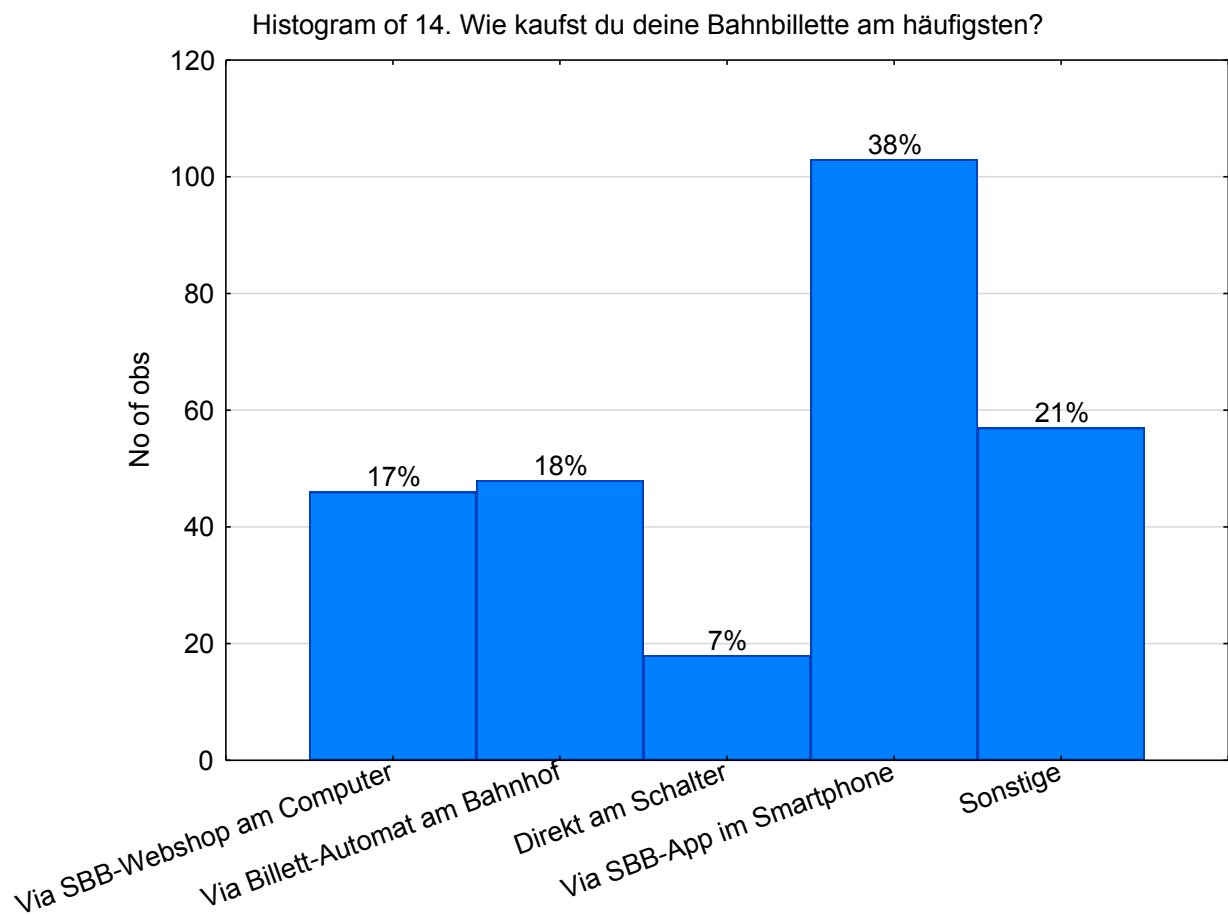
Für die Variable: „Wie oft fährst du mit der SBB?“

Category	Wie oft fährst du mit der SBB?			
	Count	Cumulative Count	Percent	Cumulative Percent
(fast) nie	33	33	12,13235	12,1324
mehrmals pro Monat	59	92	21,69118	33,8235
mehrmals pro Woche	47	139	17,27941	51,1029
(fast) täglich	93	232	34,19118	85,2941
mehrmals pro Jahr	40	272	14,70588	100,0000
Missing	0	272	0,00000	100,0000



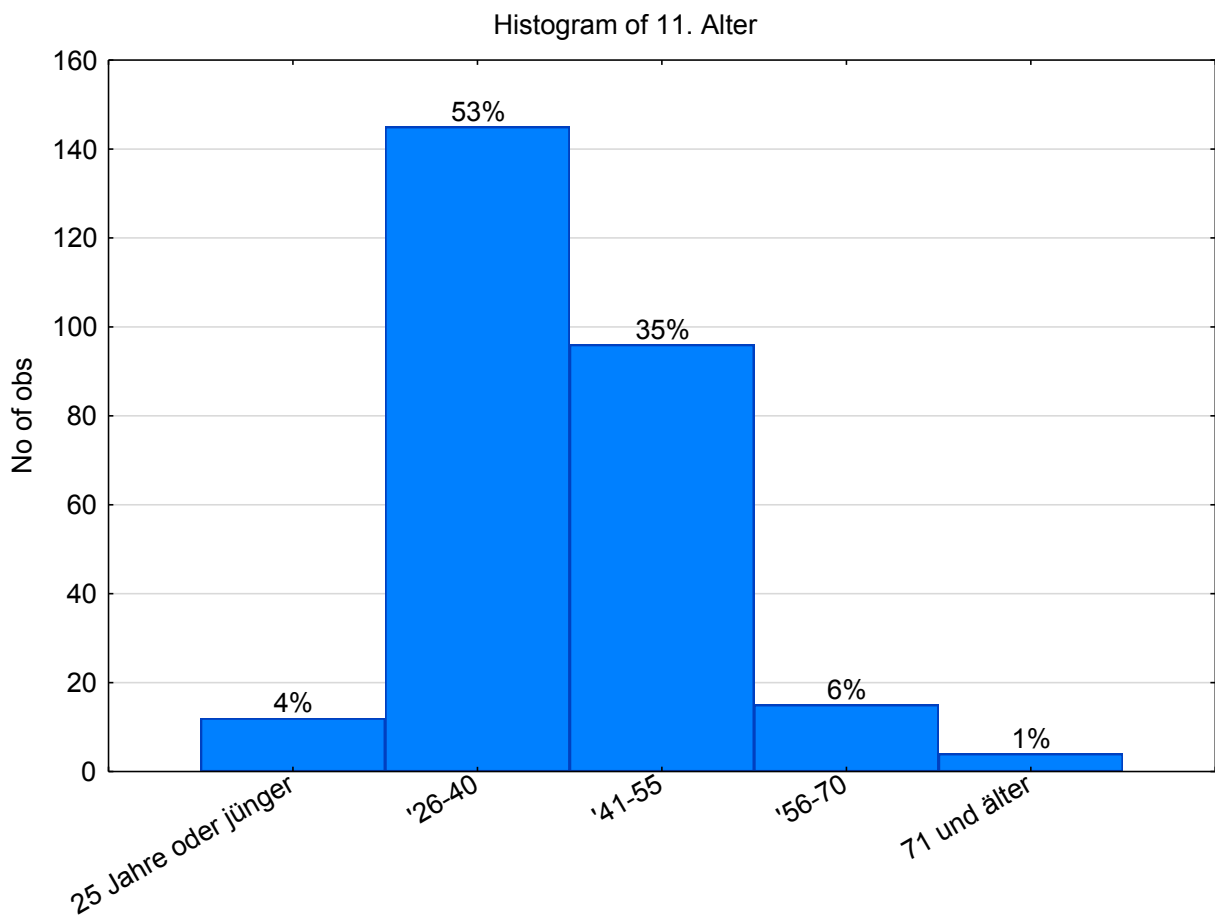
Für die Variable: „Wie kaufst du dein Bahnticket üblicherweise?“

Category	Wie kaufst du deine Bahnbillette am häufigsten?			
	Count	Cumulative Count	Percent	Cumulative Percent
Via SBB-Webshop am Computer	46	46	16,91176	16,9118
Sonstige	57	103	20,95588	37,8676
Direkt am Schalter	18	121	6,61765	44,4853
Via SBB-App im Smartphone	103	224	37,86765	82,3529
Via Billett-Automat am Bahnhof	48	272	17,64706	100,0000
Missing	0	272	0,00000	100,0000



Für die Variable: Altersgruppe

Category	Alter			
	Count	Cumulative Count	Percent	Cumulative Percent
25 Jahre oder jünger	12	12	4,41176	4,4118
'26-40	145	157	53,30882	57,7206
'41-55	96	253	35,29412	93,0147
'56-70	15	268	5,51471	98,5294
71 und älter	4	272	1,47059	100,0000
Missing	0	272	0,00000	100,0000



## Mittelwerte und Mediane

Variable	Descriptive Statistics (Datensatz: 3 f		
	Valid N	Mean	Median
langsam (1) - schnell (7)	261	2,727969	2,000000
nutzlos (1) - nützlich (7)	255	4,062745	4,000000
verwirrend (1) - übersichtlich (7)	254	4,350394	4,000000
unattraktiv (1) - attraktiv (7)	254	3,783465	4,000000
unberechenbar (1) - voraussagbar (7)	253	4,470356	4,000000
konventionell (1) - originell (7)	253	3,462451	4,000000
lästig (1) - angenehm (7)	253	3,837945	4,000000
unsympathisch (1) - sympathisch (7)	254	4,000000	4,000000
langweilig (1) - spannend (7)	254	3,255906	4,000000
unverständlich (1) - verständlich (7)	255	4,976471	6,000000

Variable	PROTOTYP=1 Descriptive Statistics (Spreadsheet i		
	Valid N	Mean	Median
langsam (1) - schnell (7)	58	2,310345	2,000000
nutzlos (1) - nützlich (7)	52	3,076923	3,000000
verwirrend (1) - übersichtlich (7)	51	2,960784	3,000000
unattraktiv (1) - attraktiv (7)	51	2,686275	3,000000
unberechenbar (1) - voraussagbar (7)	50	3,460000	4,000000
konventionell (1) - originell (7)	50	3,020000	3,500000
lästig (1) - angenehm (7)	50	3,220000	4,000000
unsympathisch (1) - sympathisch (7)	51	3,078431	4,000000
langweilig (1) - spannend (7)	51	2,764706	3,000000
unverständlich (1) - verständlich (7)	52	3,076923	4,000000

Variable	PROTOTYP=2 Descriptive Statistics (Spreadsheet i		
	Valid N	Mean	Median
langsam (1) - schnell (7)	69	2,550725	2,000000
nutzlos (1) - nützlich (7)	69	4,289855	5,000000
verwirrend (1) - übersichtlich (7)	69	4,782609	6,000000
unattraktiv (1) - attraktiv (7)	69	3,405797	4,000000
unberechenbar (1) - voraussagbar (7)	69	5,420290	6,000000
konventionell (1) - originell (7)	69	2,492754	2,000000
lästig (1) - angenehm (7)	69	3,623188	4,000000
unsympathisch (1) - sympathisch (7)	69	3,753623	4,000000
langweilig (1) - spannend (7)	69	2,681159	3,000000
unverständlich (1) - verständlich (7)	69	5,536232	6,000000

Variable	PROTOTYP=3 Descriptive Statistics (Spreadsheet		
	Valid N	Mean	Median
langsam (1) - schnell (7)	66	2,924242	3,000000
nutzlos (1) - nützlich (7)	66	4,424242	5,000000
verwirrend (1) - übersichtlich (7)	66	4,606061	5,000000
unattraktiv (1) - attraktiv (7)	66	3,818182	4,000000
unberechenbar (1) - voraussagbar (7)	66	4,484848	5,000000
konventionell (1) - originell (7)	66	2,893939	2,500000
lästig (1) - angenehm (7)	66	3,772727	4,000000
unsympathisch (1) - sympathisch (7)	66	3,924242	4,000000
langweilig (1) - spannend (7)	66	3,136364	3,000000
unverständlich (1) - verständlich (7)	66	5,348485	6,000000

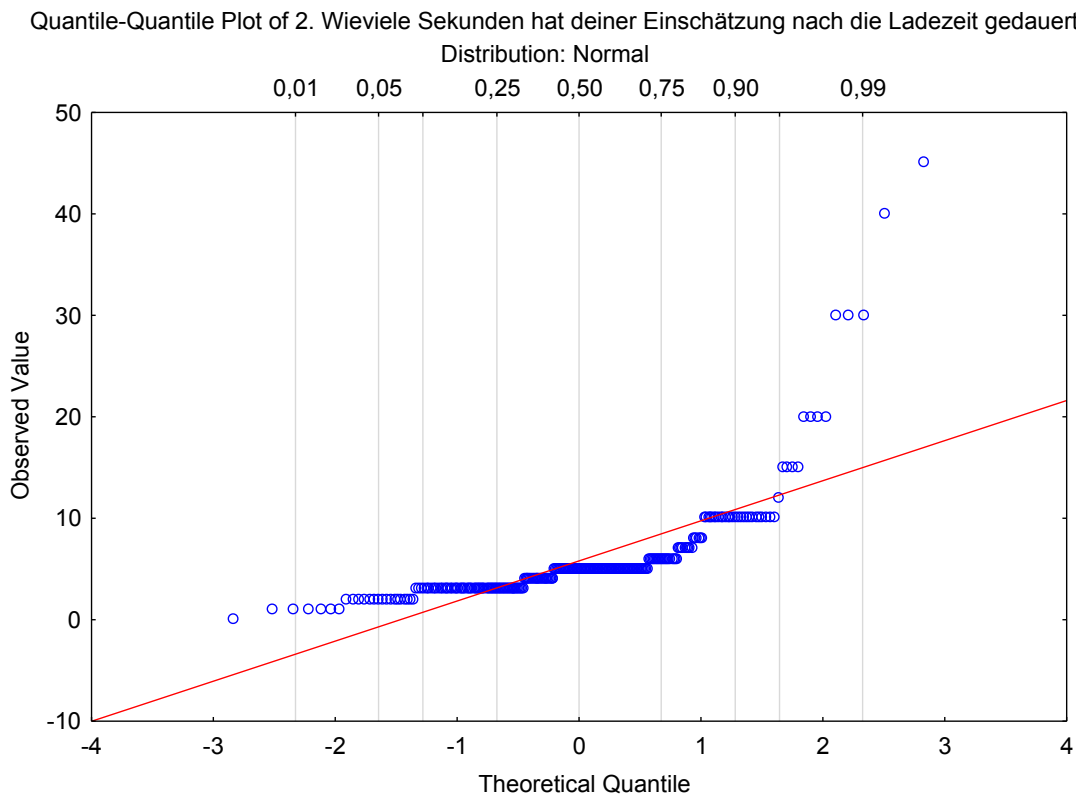
Variable	PROTOTYP=4 Descriptive Statistics (Spreadsheet ir		
	Valid N	Mean	Median
langsam (1) - schnell (7)	68	3,073529	3,000000
nutzlos (1) - nützlich (7)	68	4,235294	4,000000
verwirrend (1) - übersichtlich (7)	68	4,705882	5,000000
unattraktiv (1) - attraktiv (7)	68	4,955882	5,000000
unberechenbar (1) - voraussagbar (7)	68	4,235294	4,000000
konventionell (1) - originell (7)	68	5,323529	6,000000
lästig (1) - angenehm (7)	68	4,573529	5,000000
unsympathisch (1) - sympathisch (7)	68	5,014706	5,000000
langweilig (1) - spannend (7)	68	4,323529	4,000000
unverständlich (1) - verständlich (7)	68	5,500000	6,000000

# Validierung der Hypothesen

## Hypothese 1: Wahrnehmung der Ladedauer

- Wie viele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?
- Ich empfand die Dauer als angemessen.
- Schnell/Langsam

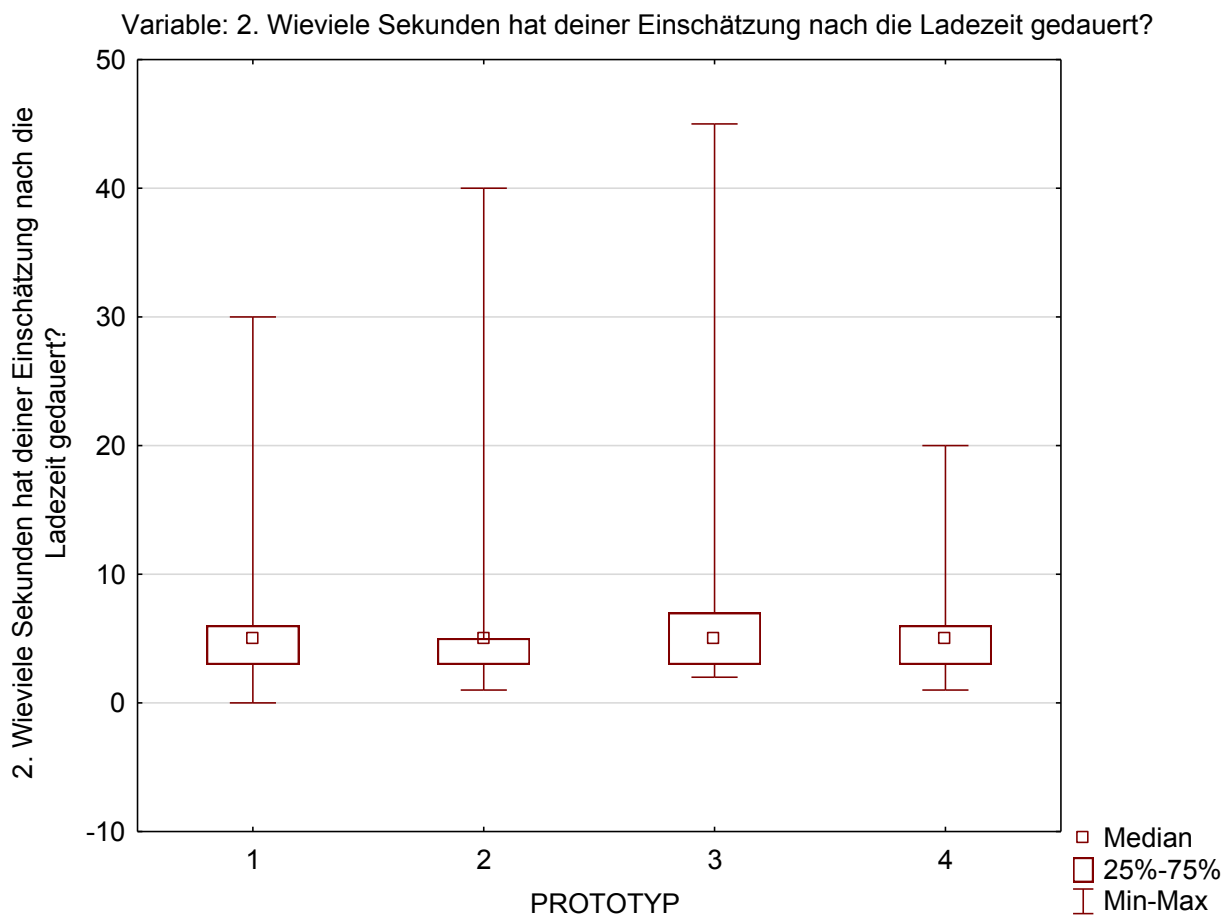
## Wie viele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?



Variable entspricht keiner Normalverteilung

- Kruskal-Wallis-Rangsummentest
- P-Wert 0.6994, nicht signifikant → kein Unterschied zwischen den Prototypen in der zentralen Tendenz der geschätzten Dauer

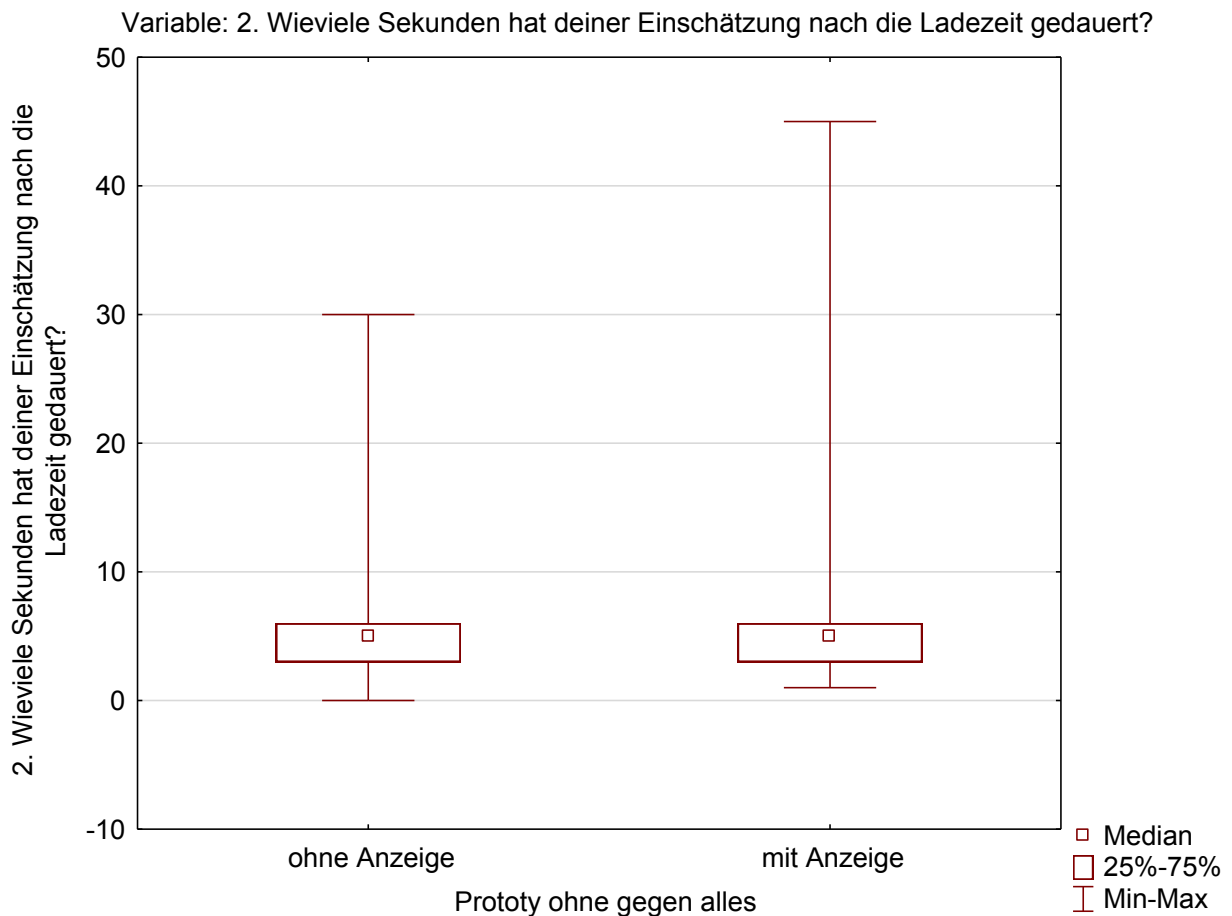
Depend.:		Kruskal-Wallis ANOVA by Ranks; 2. Wieviele Sekunden hat dein			
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?		Independent (grouping) variable: PROTOTYP			
		Kruskal-Wallis test: H ( 3, N= 269) = 1,426397 p = ,6994			
	Code	Valid N	Sum of Ranks	Mean Rank	
1	1	66	9166,000	138,8788	
2	2	69	8721,500	126,3986	
3	3	66	9286,500	140,7045	
4	4	68	9141,000	134,4265	



Gruppe 1 vs. Alle anderen (Anzeige vs. Ohne Anzeige)

- Mann-Whitney-U-Test
- P-Wert 0.6344, nicht signifikant

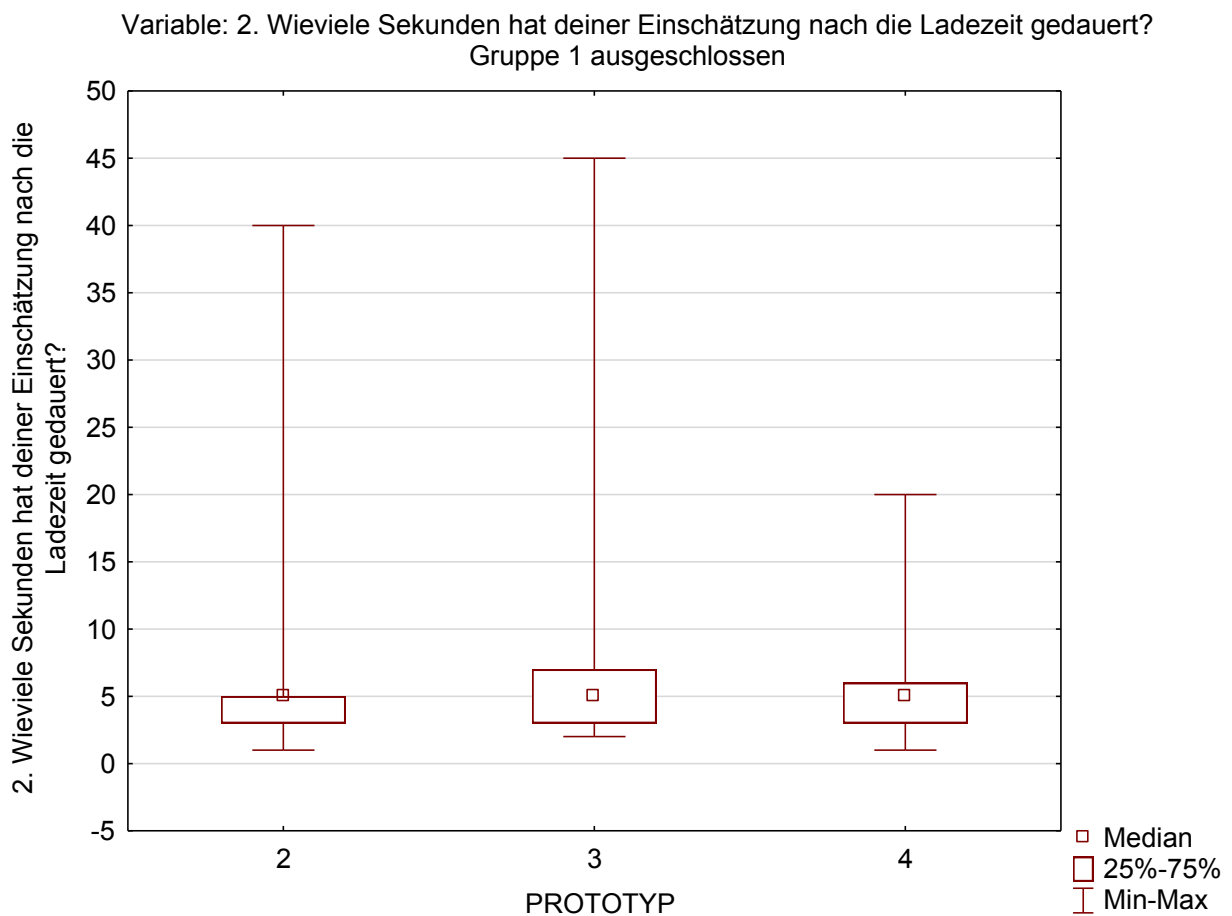
Mann-Whitney U Test (w/ continuity correction) (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_SBB_Ladeanzeige)									
By variable Prototy ohne gegen alles									
Marked tests are significant at p <.05000									
variable	Rank Sum ohne Anzeige	Rank Sum mit Anzeige	U	Z	p-value	Z adjusted	p-value	Valid N ohne Anzeige	Valid N mit Anzeige
2. Wieviele Sekunden hat deiner Einschätzung n	9166,000	27149,00	6443,000	0,465350	0,641681	0,475516	0,634419	66	203



Die 3 Gruppen mit Anzeigen untereinander

- Kruskal-Wallis-Test der Rangsummen
- P-Wert 0.5474, nicht signifikant

Depend.:		Kruskal-Wallis ANOVA by Ranks; 2. Wieviele Sekunde Independent (grouping) variable: PROTOTYP Kruskal-Wallis test: H ( 2, N= 203) =1,205240 p =,5474 Exclude condition: v6 = 1		
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Code	Valid N	Sum of Ranks	Mean Rank
2	2	69	6655,500	96,4565
3	3	66	7080,000	107,2727
4	4	68	6970,500	102,5074



## Wie viele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?

Ich empfand die Dauer als angemessen.

Korrelieren die geschätzten Dauern und die Aussagen über die Angemessenheit der Ladedauer?

### Spearman's Rangkorrelationskoeffizient – alle Prototypen

Variable	Spearman Rank Order Correlations, MD pairwise deleted Marked correlations are significant at p <,05000	
	2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Ich empfand die Dauer als angemessen.
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	1,000000	-0,093552
Ich empfand die Dauer als angemessen.	-0,093552	1,000000

### Spearman's Rangkorrelationskoeffizient – einzeln für die Prototypen

Variable	Spearman Rank Order Correlations MD pairwise deleted Marked correlations are significant at p <,05000 Include condition: v6 = 1	
	2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Ich empfand die Dauer als angemessen.
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	1,000000	-0,291705
Ich empfand die Dauer als angemessen.	-0,291705	1,000000

Variable	Spearman Rank Order Correlations MD pairwise deleted Marked correlations are significant at p <,05000 Include condition: v6 = 2	
	2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Ich empfand die Dauer als angemessen.
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	1,000000	-0,130462
Ich empfand die Dauer als angemessen.	-0,130462	1,000000

Variable	Spearman Rank Order Correlations MD pairwise deleted Marked correlations are significant at p <,05000 Include condition: v6 = 3	
	2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Ich empfand die Dauer als angemessen.
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	1,000000	0,059284
Ich empfand die Dauer als angemessen.	0,059284	1,000000

Variable	Spearman Rank Order Correlations MD pairwise deleted Marked correlations are significant at p <,05000 Include condition: v6 = 4	
	2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	Ich empfand die Dauer als angemessen.
2. Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert?	1,000000	-0,039357
Ich empfand die Dauer als angemessen.	-0,039357	1,000000

## Ich empfand die Dauer als angemessen

Alle Prototypen gegeneinander

PROTOTYP	2-Way Summary Table: Observed Frequencies Marked cells have counts > 10					Row Totals
	Ich empfand die Dauer als angemessen. 1	Ich empfand die Dauer als angemessen. 2	Ich empfand die Dauer als angemessen. 3	Ich empfand die Dauer als angemessen. 4	Ich empfand die Dauer als angemessen. 5	
1	24	26	7	7	2	66
Column %	30%	25%	18%	18%	18%	
Row %	36%	39%	11%	11%	3%	
Total %	9%	10%	3%	3%	1%	25%
2	22	27	9	9	2	69
Column %	28%	26%	24%	23%	18%	
Row %	32%	39%	13%	13%	3%	
Total %	8%	10%	3%	3%	1%	26%
3	9	30	12	13	2	66
Column %	11%	29%	32%	33%	18%	
Row %	14%	45%	18%	20%	3%	
Total %	3%	11%	4%	5%	1%	25%
4	24	19	10	10	5	68
Column %	30%	19%	26%	26%	45%	
Row %	35%	28%	15%	15%	7%	
Total %	9%	7%	4%	4%	2%	25%
Totals	79	102	38	39	11	269
Total %	29%	38%	14%	14%	4%	100%

Statistic	Statistics: PROTOTYP(4) x Ich empfand die Dauer als angemessen.(5)		
	Chi-square	df	p
Pearson Chi-square	16,27793	df=12	p=,17884
M-L Chi-square	17,32423	df=12	p=,13780

Nur die 3 Prototypen mit Anzeige gegeneinander

2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen & Marked cells have counts > 10 Exclude condition: v6 =1						
PROTOTYP	Ich empfand die Dauer als angemessen. 1	Ich empfand die Dauer als angemessen. 2	Ich empfand die Dauer als angemessen 3	Ich empfand die Dauer als angemessen. 4	Ich empfand die Dauer als angemessen 5	Row Totals
2	22	27	9	9	2	69
Column %	40%	36%	29%	28%	22%	
Row %	32%	39%	13%	13%	3%	
Total %	11%	13%	4%	4%	1%	34%
3	9	30	12	13	2	66
Column %	16%	39%	39%	41%	22%	
Row %	14%	45%	18%	20%	3%	
Total %	4%	15%	6%	6%	1%	33%
4	24	19	10	10	5	68
Column %	44%	25%	32%	31%	56%	
Row %	35%	28%	15%	15%	7%	
Total %	12%	9%	5%	5%	2%	33%
Totals	55	76	31	32	9	203
Total %	27%	37%	15%	16%	4%	100%

Statistics: PROTOTYP(3) x Ich empfand die Dauer als angemessen. Ex			
Statistic	Chi-square	df	p
Pearson Chi-square	13,09918	df=8	p=,10848
M-L Chi-square	13,83269	df=8	p=,08623

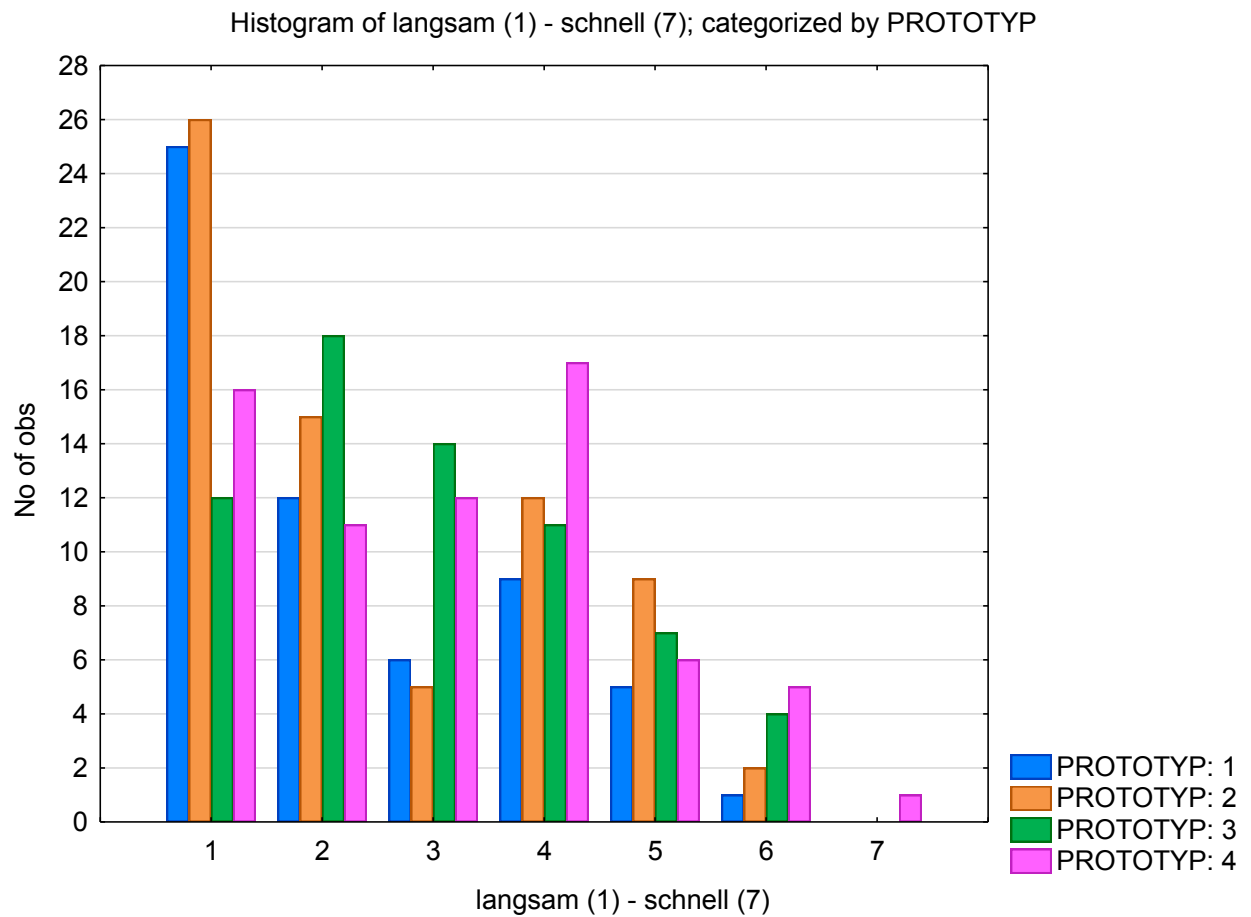
## Prototyp ohne Anzeige gegen die Prototypen mit Anzeige

Prototy ohne gegen alles	2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Marked cells have counts > 10)					Row Totals
	Ich empfand die Dauer als angemessen. 1	Ich empfand die Dauer als angemessen. 2	Ich empfand die Dauer als angemessen. 3	Ich empfand die Dauer als angemessen. 4	Ich empfand die Dauer als angemessen. 5	
ohne Anzeige	24	26	7	7	2	66
Column %	30%	25%	18%	18%	18%	
Row %	36%	39%	11%	11%	3%	
Total %	9%	10%	3%	3%	1%	25%
mit Anzeige	55	76	31	32	9	203
Column %	70%	75%	82%	82%	82%	
Row %	27%	37%	15%	16%	4%	
Total %	20%	28%	12%	12%	3%	75%
Totals	79	102	38	39	11	269
Total %	29%	38%	14%	14%	4%	100%

Statistic	Statistics: Prototy ohne gegen alles(2)		
	Chi-square	df	p
Pearson Chi-square	3,428490	df=4	p=,48883
M-L Chi-square	3,491543	df=4	p=,47917

## Langsam – schnell

Statistic	Statistics: PROTOTYP(4) x langsam (1)		
	Chi-square	df	p
Pearson Chi-square	24,88836	df=18	p=,12802
M-L Chi-square	25,19764	df=18	p=,11957



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Marked cells have counts > 10								
PROTOTYP	langsam (1) · schnell (7) 1	langsam (1) · - schnell (7) 2	langsam (1) · - schnell (7) 3	langsam (1) · - schnell (7) 4	langsam (1) · - schnell (7) 5	langsam (1) · - schnell (7) 6	langsam (1) · - schnell (7) 7	Row Totals
1	25	12	6	9	5	1	0	58
Column %	32%	21%	16%	18%	19%	8%	0%	
Row %	43%	21%	10%	16%	9%	2%	0%	
Total %	10%	5%	2%	3%	2%	0%	0%	22%
2	26	15	5	12	9	2	0	69
Column %	33%	27%	14%	24%	33%	17%	0%	
Row %	38%	22%	7%	17%	13%	3%	0%	
Total %	10%	6%	2%	5%	3%	1%	0%	26%
3	12	18	14	11	7	4	0	66
Column %	15%	32%	38%	22%	26%	33%	0%	
Row %	18%	27%	21%	17%	11%	6%	0%	
Total %	5%	7%	5%	4%	3%	2%	0%	25%
4	16	11	12	17	6	5	1	68
Column %	20%	20%	32%	35%	22%	42%	100%	
Row %	24%	16%	18%	25%	9%	7%	1%	
Total %	6%	4%	5%	7%	2%	2%	0%	26%
Totals	79	56	37	49	27	12	1	261
Total %	30%	21%	14%	19%	10%	5%	0%	100%

## Hypothese 2: Visuelle Ästhetik

- Unattraktiv – attraktiv
- Konventionell – originell
- Unsympathisch – sympathisch
- Langweilig – spannend
- Ich empfand die Darstellung als optisch attraktiv

### Unattraktiv – attraktiv

Omnibus – mit Prototyp 2, 3 und 4

Statistic	Statistics: PROTOTYP(3) x unattrakti		
	Exclude condition: v6 = 1		
	Chi-square	df	p
Pearson Chi-square	42,54392	df=12	p=,00003
M-L Chi-square	42,77929	df=12	p=,00002

Posthoc-Tests inklusive Bonferroni-Holms-Korrektur  
Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x unattrakti		
	Chi-square	df	p
Pearson Chi-square	8,128638	df=6	p=,22883
M-L Chi-square	8,294484	df=6	p=,21731

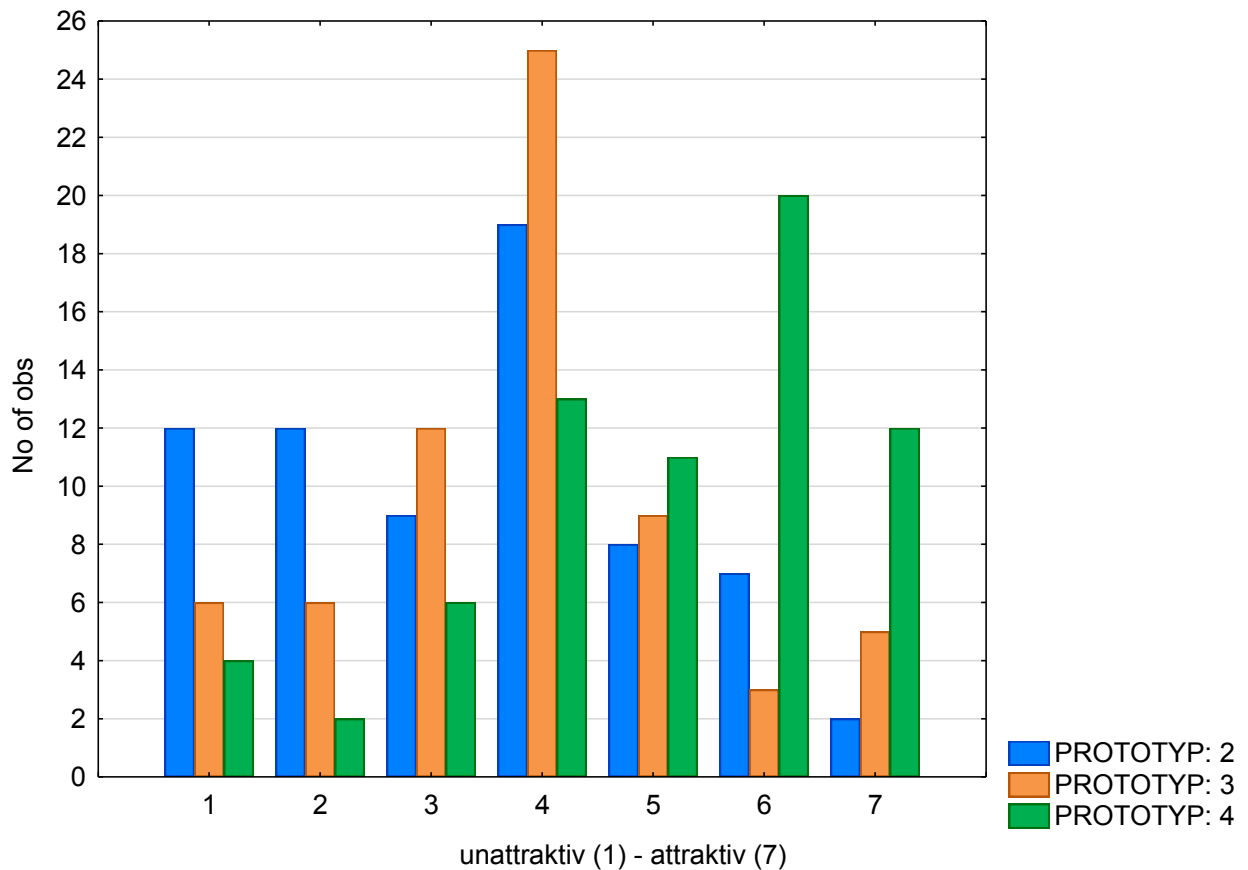
Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x unattraktiv (1) -		
	Chi-square	df	p
Pearson Chi-square	26,73778	df=6	p=,00016 (0.00048)
M-L Chi-square	28,76663	df=6	p=,00007 (0.00021)

Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x unattraktiv (1) - a		
	Chi-square	df	p
Pearson Chi-square	23,81250	df=6	p=,00057 (0.00114)
M-L Chi-square	25,60298	df=6	p=,00026 (0.00052)

Histogram of unattraktiv (1) - attraktiv (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_)								
Marked cells have counts > 10								
Exclude condition: v6 = 1								
PROTOTYP	unattraktiv (1) - attraktiv (7) 1	unattraktiv (1) - attraktiv (7) 2	unattraktiv (1) - attraktiv (7) 3	unattraktiv (1) - attraktiv (7) 4	unattraktiv (1) - attraktiv (7) 5	unattraktiv (1) - attraktiv (7) 6	unattraktiv (1) - attraktiv (7) 7	Row Totals
2	12	12	9	19	8	7	2	69
Column %	55%	60%	33%	33%	29%	23%	11%	
Row %	17%	17%	13%	28%	12%	10%	3%	
Total %	6%	6%	4%	9%	4%	3%	1%	34%
3	6	6	12	25	9	3	5	66
Column %	27%	30%	44%	44%	32%	10%	26%	
Row %	9%	9%	18%	38%	14%	5%	8%	
Total %	3%	3%	6%	12%	4%	1%	2%	33%
4	4	2	6	13	11	20	12	68
Column %	18%	10%	22%	23%	39%	67%	63%	
Row %	6%	3%	9%	19%	16%	29%	18%	
Total %	2%	1%	3%	6%	5%	10%	6%	33%
Totals	22	20	27	57	28	30	19	203
Total %	11%	10%	13%	28%	14%	15%	9%	100%

## Konventionell – originell

Omnibus – mit Prototyp 2, 3 und 4

Statistics: PROTOTYP(3) x konventionell Exclude condition: v6 = 1			
Statistic	Chi-square	df	p
Pearson Chi-square	100,1723	df=12	p=,00000
M-L Chi-square	106,9917	df=12	p=0,0000

Posthoc-Tests inklusive Bonferroni-Holms-Korrektur  
Prototyp 2-3

Statistics: PROTOTYP(2) x konvention			
Statistic	Chi-square	df	p
Pearson Chi-square	8,806010	df=6	p=,18479
M-L Chi-square	8,977735	df=6	p=,17483

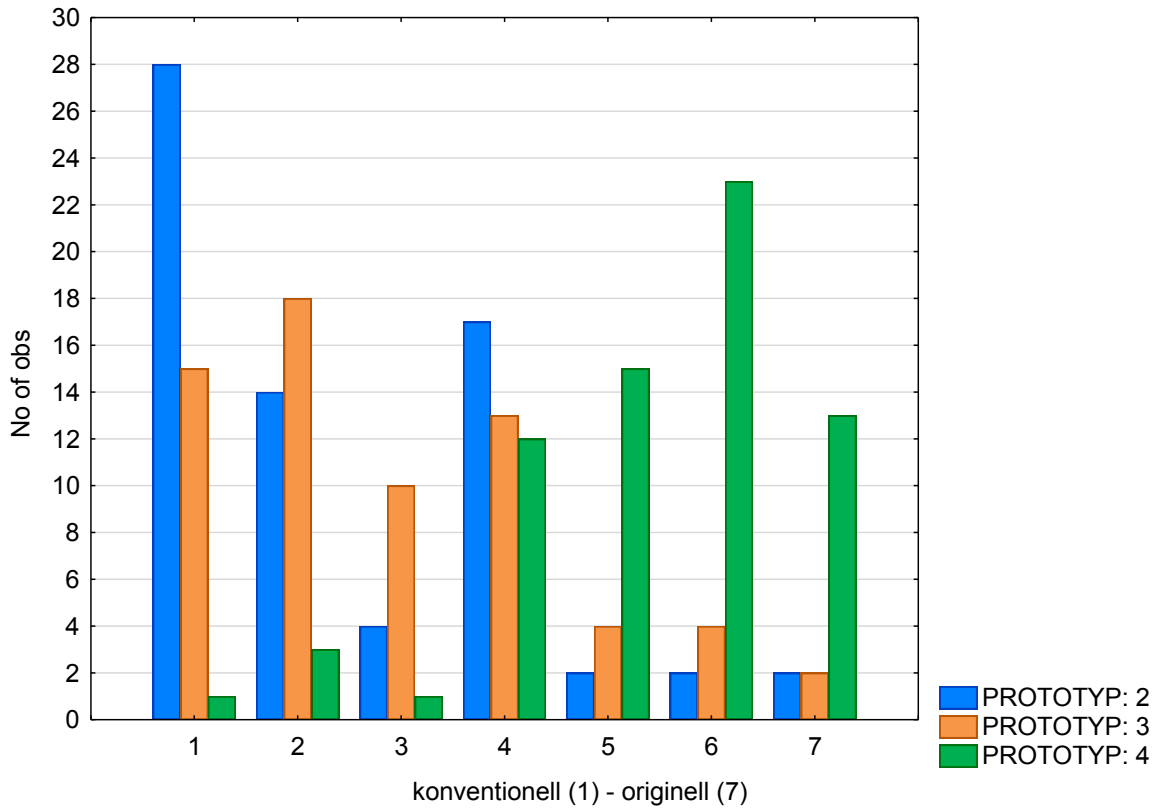
Prototyp 2-4

Statistics: PROTOTYP(2) x konvention			
Statistic	Chi-square	df	p
Pearson Chi-square	70,56195	df=6	p=,00000
M-L Chi-square	82,99731	df=6	p=,00000

Prototyp 3-4

Statistics: PROTOTYP(2) x konvention			
Statistic	Chi-square	df	p
Pearson Chi-square	58,15649	df=6	p=,00000
M-L Chi-square	65,71887	df=6	p=,00000

Histogram of konventionell (1) - originell (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung\_SBB\_Ladeanzeige)  
 Marked cells have counts > 10  
 Exclude condition: v6 = 1

PROTOTYP	konventionell (1) - originell (7) 1	konventionell (1) - originell (7) 2	konventionell (1) - originell (7) 3	konventionell (1) - originell (7) 4	konventionell (1) - originell (7) 5	konventionell (1) - originell (7) 6	konventionell (1) - originell (7) 7	Row Totals
2	28	14	4	17	2	2	2	69
Column %	64%	40%	27%	40%	10%	7%	12%	
Row %	41%	20%	6%	25%	3%	3%	3%	
Total %	14%	7%	2%	8%	1%	1%	1%	34%
3	15	18	10	13	4	4	2	66
Column %	34%	51%	67%	31%	19%	14%	12%	
Row %	23%	27%	15%	20%	6%	6%	3%	
Total %	7%	9%	5%	6%	2%	2%	1%	33%
4	1	3	1	12	15	23	13	68
Column %	2%	9%	7%	29%	71%	79%	76%	
Row %	1%	4%	1%	18%	22%	34%	19%	
Total %	0%	1%	0%	6%	7%	11%	6%	33%
Totals	44	35	15	42	21	29	17	203
Total %	22%	17%	7%	21%	10%	14%	8%	100%

## Unsympathisch – sympathisch

Omnibus – mit Prototyp 2, 3 und 4

Statistics: PROTOTYP(3) x unsympathisch			
Exclude condition: v6 = 1			
Statistic	Chi-square	df	p
Pearson Chi-square	45,61855	df=12	p=,00001
M-L Chi-square	46,68249	df=12	p=,00001

Posthoc-Tests inklusive Bonferroni-Holms-Korrektur  
Prototyp 2-3

Statistics: PROTOTYP(2) x unsympathisch			
Statistic	Chi-square	df	p
Pearson Chi-square	1,841885	df=6	p=,93365
M-L Chi-square	1,867664	df=6	p=,93146

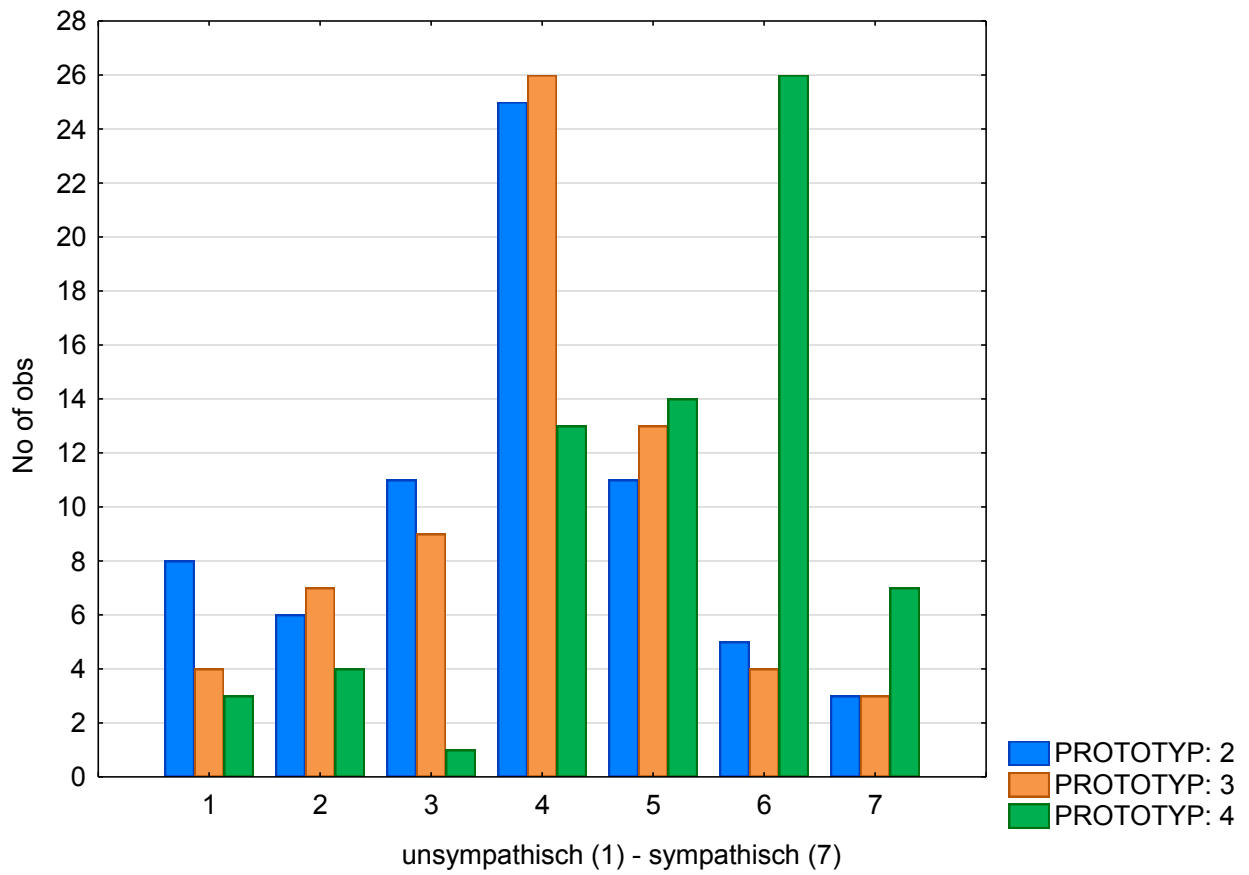
Prototyp 2-4

Statistics: PROTOTYP(2) x unsympathisch (1) - symp			
Statistic	Chi-square	df	p
Pearson Chi-square	30,97569	df=6	p=,00003 (0,00009)
M-L Chi-square	33,95012	df=6	p=,00001 (0,00003)

Prototyp 3-4

Statistics: PROTOTYP(2) x unsympathisch (1) - s			
Statistic	Chi-square	df	p
Pearson Chi-square	29,44145	df=6	p=,00005 (0,0001)
M-L Chi-square	32,43186	df=6	p=,00001 (0,00002)

Histogram of unsympathisch (1) - sympathisch (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung\_SBB\_Ladeanzeige)  
 Marked cells have counts > 10  
 Exclude condition: v6 = 1

PROTOTYP	unsympathisch (1) - sympathisch (7) 1	unsympathisch (1) - sympathisch (7) 2	unsympathisch (1) - sympathisch (7) 3	unsympathisch (1) - sympathisch (7) 4	unsympathisch (1) - sympathisch (7) 5	unsympathisch (1) - sympathisch (7) 6	unsympathisch (1) - sympathisch (7) 7	Row Totals
2	8	6	11	25	11	5	3	69
Column %	53%	35%	52%	39%	29%	14%	23%	
Row %	12%	9%	16%	36%	16%	7%	4%	
Total %	4%	3%	5%	12%	5%	2%	1%	34%
3	4	7	9	26	13	4	3	66
Column %	27%	41%	43%	41%	34%	11%	23%	
Row %	6%	11%	14%	39%	20%	6%	5%	
Total %	2%	3%	4%	13%	6%	2%	1%	33%
4	3	4	1	13	14	26	7	68
Column %	20%	24%	5%	20%	37%	74%	54%	
Row %	4%	6%	1%	19%	21%	38%	10%	
Total %	1%	2%	0%	6%	7%	13%	3%	33%
Totals	15	17	21	64	38	35	13	203
Total %	7%	8%	10%	32%	19%	17%	6%	100%

## Langweilig – spannend

Omnibus – mit Prototyp 2, 3 und 4

Statistics: PROTOTYP(3) x langweilig			
Exclude condition: v6 = 1			
Statistic	Chi-square	df	p
Pearson Chi-square	51,13140	df=12	p=,00000
M-L Chi-square	53,67589	df=12	p=,00000

Posthoc-Tests inklusive Bonferroni-Holms-Korrektur  
Prototyp 2-3

Statistics: PROTOTYP(2) x langweilig			
Statistic	Chi-square	df	p
Pearson Chi-square	5,379789	df=6	p=,49610
M-L Chi-square	5,800457	df=6	p=,44591

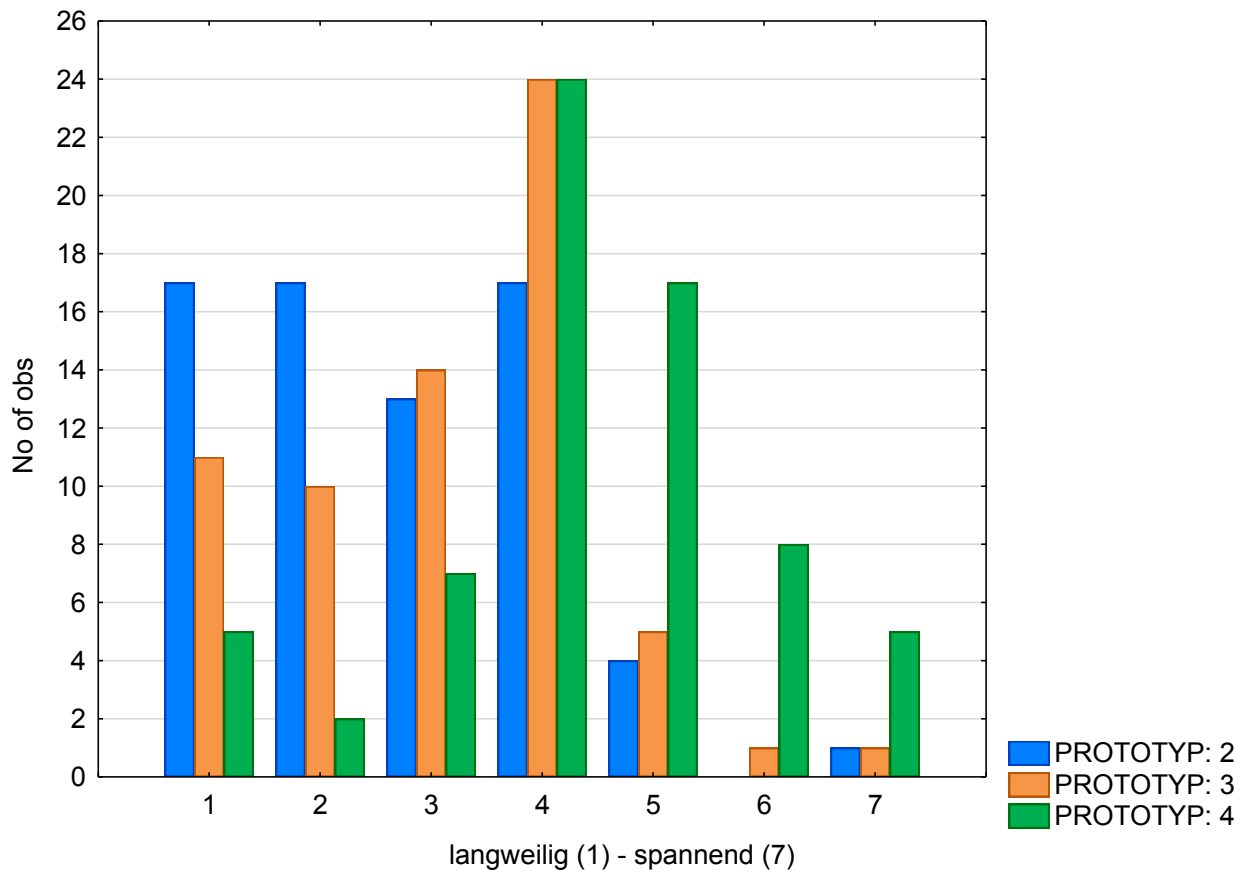
Prototyp 2-4

Statistics: PROTOTYP(2) x langweilig			
Statistic	Chi-square	df	p
Pearson Chi-square	40,09180	df=6	p=,00000
M-L Chi-square	46,15395	df=6	p=,00000

Prototyp 3-4

Statistics: PROTOTYP(2) x langweilig (1) - span			
Statistic	Chi-square	df	p
Pearson Chi-square	24,54885	df=6	p=,00041 (0,00082)
M-L Chi-square	26,50169	df=6	p=,00018 (0,00036)

Histogram of langweilig (1) - spannend (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_SBB_Ladeanz Marked cells have counts > 10 Exclude condition: v6 = 1								
PROTOTYP	langweilig (1) - spannend (7) 1	langweilig (1) - spannend (7) 2	langweilig (1) - spannend (7) 3	langweilig (1) - spannend (7) 4	langweilig (1) - spannend (7) 5	langweilig (1) - spannend (7) 6	langweilig (1) - spannend (7) 7	Row Totals
2	17	17	13	17	4	0	1	69
Column %	52%	59%	38%	26%	15%	0%	14%	
Row %	25%	25%	19%	25%	6%	0%	1%	
Total %	8%	8%	6%	8%	2%	0%	0%	34%
3	11	10	14	24	5	1	1	66
Column %	33%	34%	41%	37%	19%	11%	14%	
Row %	17%	15%	21%	36%	8%	2%	2%	
Total %	5%	5%	7%	12%	2%	0%	0%	33%
4	5	2	7	24	17	8	5	68
Column %	15%	7%	21%	37%	65%	89%	71%	
Row %	7%	3%	10%	35%	25%	12%	7%	
Total %	2%	1%	3%	12%	8%	4%	2%	33%
Totals	33	29	34	65	26	9	7	203
Total %	16%	14%	17%	32%	13%	4%	3%	100%

## Ich empfand die Darstellung als optisch attraktiv

Omnibus – mit Prototyp 2, 3 und 4

Statistics: PROTOTYP(3) x Ich empfar			
Exclude condition: v6 = 1			
Statistic	Chi-square	df	p
Pearson Chi-square	38,54168	df=8	p=,00001
M-L Chi-square	39,54151	df=8	p=,00000

Posthoc-Tests inklusive Bonferroni-Holms-Korrektur  
Prototyp 2-3

Statistics: PROTOTYP(2) x Ich empfar			
Statistic	Chi-square	df	p
Pearson Chi-square	4,358455	df=4	p=,35966
M-L Chi-square	4,405366	df=4	p=,35392

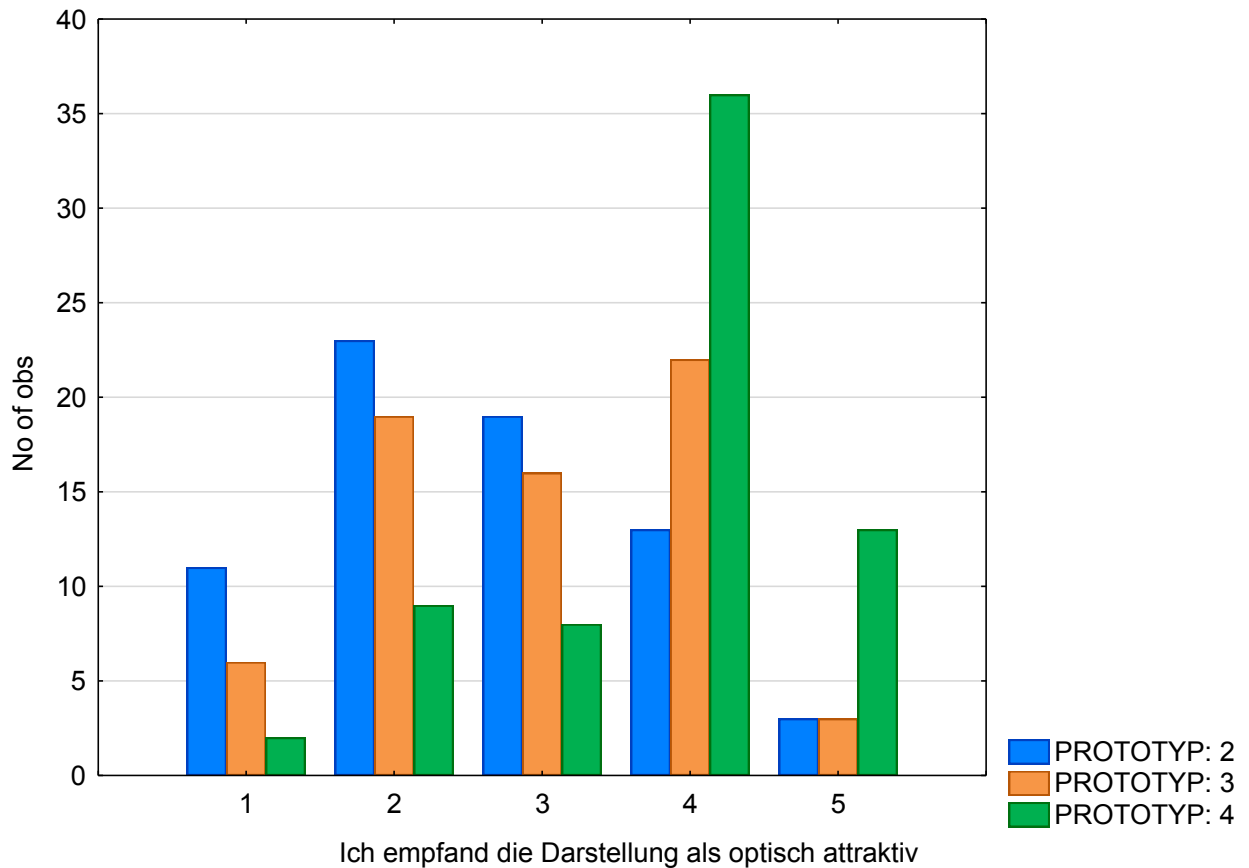
Prototyp 2-4

Statistics: PROTOTYP(2) x Ich empfar			
Statistic	Chi-square	df	p
Pearson Chi-square	33,87767	df=4	p=,00000
M-L Chi-square	35,77404	df=4	p=,00000

Prototyp 3-4

Statistics: PROTOTYP(2) x Ich empfand die Dar			
Statistic	Chi-square	df	p
Pearson Chi-square	17,84153	df=4	p=,00133 (0,00266)
M-L Chi-square	18,58418	df=4	p=,00095 (0,0019)

Histogram of Ich empfand die Darstellung als optisch attraktiv; categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2  
 Marked cells have counts > 10  
 Exclude condition: v6 = 1

PROTOTYP	Ich empfand die Darstellung als optisch attraktiv 1	Ich empfand die Darstellung als optisch attraktiv 2	Ich empfand die Darstellung als optisch attraktiv 3	Ich empfand die Darstellung als optisch attraktiv 4	Ich empfand die Darstellung als optisch attraktiv 5	Row Totals
2	11	23	19	13	3	69
Column %	58%	45%	44%	18%	16%	
Row %	16%	33%	28%	19%	4%	
Total %	5%	11%	9%	6%	1%	34%
3	6	19	16	22	3	66
Column %	32%	37%	37%	31%	16%	
Row %	9%	29%	24%	33%	5%	
Total %	3%	9%	8%	11%	1%	33%
4	2	9	8	36	13	68
Column %	11%	18%	19%	51%	68%	
Row %	3%	13%	12%	53%	19%	
Total %	1%	4%	4%	18%	6%	33%
Totals	19	51	43	71	19	203
Total %	9%	25%	21%	35%	9%	100%

### Hypothese 3: Benutzerzufriedenheit mit der Ladedauer

- Nutzlos – nützlich
- Verwirrend – übersichtlich
- Lästig – angenehm
- Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen

#### Nutzlos – nützlich

##### Omnibus

Statistic	Statistics: PROTOTYP(4) x nutzlos (1)		
	Chi-square	df	p
Pearson Chi-square	33,81875	df=18	p=,01326
M-L Chi-square	35,10303	df=18	p=,00917

##### Posthoc-Tests inklusive Bonferroni-Holms-Korrektur Prototyp 1-2

Statistic	Statistics: PROTOTYP(2) x nutzlos (1) - nützlich		
	Chi-square	df	p
Pearson Chi-square	18,33126	df=6	p=,00546 (0,0327)
M-L Chi-square	20,35901	df=6	p=,00239 (0,0143)

##### Prototyp 1-3

Statistic	Statistics: PROTOTYP(2) x nutzlos (1) - nützlich (7)		
	Chi-square	df	p
Pearson Chi-square	17,06016	df=6	p=,00906 (0,0453)
M-L Chi-square	18,32871	df=6	p=,00546 (0,0273)

##### Prototyp 1-4

Statistic	Statistics: PROTOTYP(2) x nutzlos (1) - nützlich (7)		
	Chi-square	df	p
Pearson Chi-square	14,34906	df=6	p=,02597 (0,104)
M-L Chi-square	14,75816	df=6	p=,02222 (0,0888)

##### Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x nutzlos (1)		
	Chi-square	df	p
Pearson Chi-square	3,025304	df=6	p=,80566
M-L Chi-square	3,046373	df=6	p=,80300

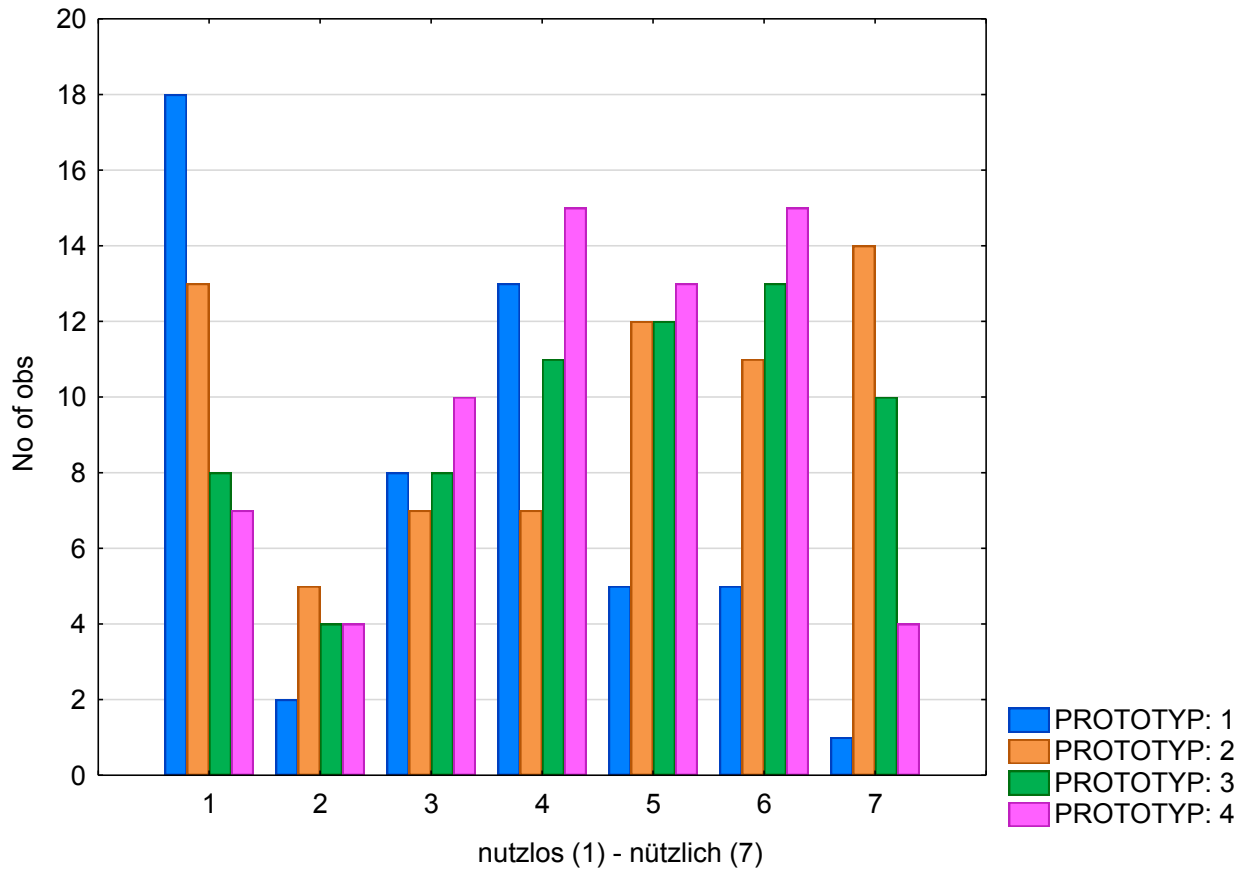
##### Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x nutzlos (1)		
	Chi-square	df	p
Pearson Chi-square	11,55387	df=6	p=,07269
M-L Chi-square	11,98283	df=6	p=,06235

##### Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x nutzlos (1)		
	Chi-square	df	p
Pearson Chi-square	3,629517	df=6	p=,72667
M-L Chi-square	3,716944	df=6	p=,71492

Histogram of nutzlos (1) - nützlich (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in A)  
Marked cells have counts > 10

PROTOTYP	nutzlos (1) - nützlich (7) 1	nutzlos (1) - nützlich (7) 2	nutzlos (1) - nützlich (7) 3	nutzlos (1) - nützlich (7) 4	nutzlos (1) - nützlich (7) 5	nutzlos (1) - nützlich (7) 6	nutzlos (1) - nützlich (7) 7	Row Totals
1	18	2	8	13	5	5	1	52
Column %	39%	13%	24%	28%	12%	11%	3%	
Row %	35%	4%	15%	25%	10%	10%	2%	
Total %	7%	1%	3%	5%	2%	2%	0%	20%
2	13	5	7	7	12	11	14	69
Column %	28%	33%	21%	15%	29%	25%	48%	
Row %	19%	7%	10%	10%	17%	16%	20%	
Total %	5%	2%	3%	3%	5%	4%	5%	27%
3	8	4	8	11	12	13	10	66
Column %	17%	27%	24%	24%	29%	30%	34%	
Row %	12%	6%	12%	17%	18%	20%	15%	
Total %	3%	2%	3%	4%	5%	5%	4%	26%
4	7	4	10	15	13	15	4	68
Column %	15%	27%	30%	33%	31%	34%	14%	
Row %	10%	6%	15%	22%	19%	22%	6%	
Total %	3%	2%	4%	6%	5%	6%	2%	27%
Totals	46	15	33	46	42	44	29	255
Total %	18%	6%	13%	18%	16%	17%	11%	100%

## Verwirrend – übersichtlich

### Omnibus

Statistic	Statistics: PROTOTYP(4) x verwirrend (1) - übersichtlich (7)(7)		
	Chi-square	df	p
Pearson Chi-square	58,12867	df=18	p=,00000
M-L Chi-square	64,04730	df=18	p=,00000

### Posthoc-Tests inklusive Bonferroni-Holms-Korrektur Prototyp 1-2

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	34,72349	df=6	p=,00000
M-L Chi-square	41,25706	df=6	p=,00000

### Prototyp 1-3

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	27,12273	df=6	p=,00014 (0,0005)
M-L Chi-square	32,38195	df=6	p=,00001 (0,0000)

### Prototyp 1-4

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	30,72893	df=6	p=,0003 (0,00015)
M-L Chi-square	35,90265	df=6	p=,00000

### Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	12,86623	df=6	p=,04521 (0,1356)
M-L Chi-square	13,18410	df=6	p=,04020 (0,1206)

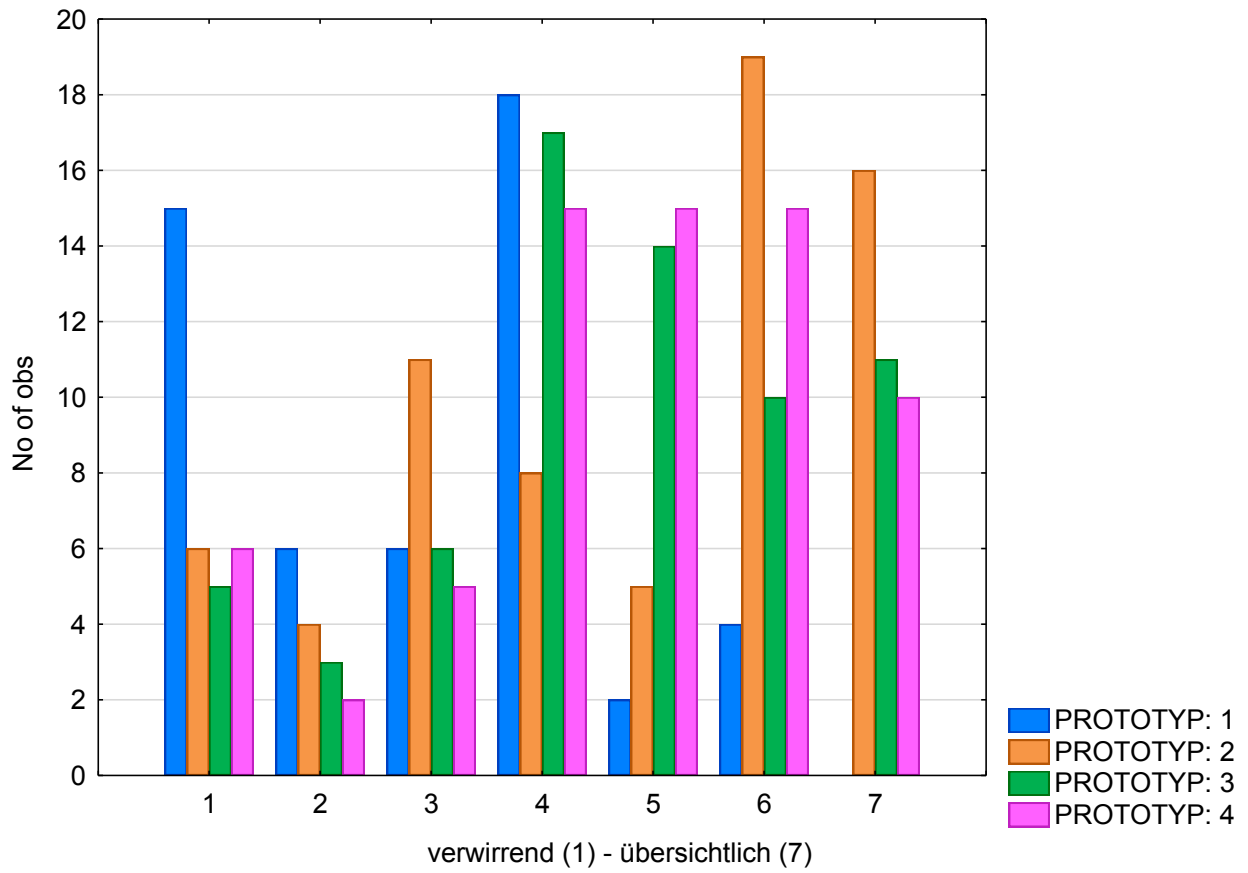
### Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	11,89564	df=6	p=,06434
M-L Chi-square	12,24419	df=6	p=,05674

### Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x verwirrend		
	Chi-square	df	p
Pearson Chi-square	1,559417	df=6	p=,95545
M-L Chi-square	1,567557	df=6	p=,95488

Histogram of verwirrend (1) - übersichtlich (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung\_SBB\_Ladeanzeige)  
Marked cells have counts > 10

PROTOTYP	verwirrend (1) - übersichtlich (7) 1	verwirrend (1) - übersichtlich (7) 2	verwirrend (1) - übersichtlich (7) 3	verwirrend (1) - übersichtlich (7) 4	verwirrend (1) - übersichtlich (7) 5	verwirrend (1) - übersichtlich (7) 6	verwirrend (1) - übersichtlich (7) 7	Row Totals
1	15	6	6	18	2	4	0	51
Column %	47%	40%	21%	31%	6%	8%	0%	
Row %	29%	12%	12%	35%	4%	8%	0%	
Total %	6%	2%	2%	7%	1%	2%	0%	20%
2	6	4	11	8	5	19	16	69
Column %	19%	27%	39%	14%	14%	40%	43%	
Row %	9%	6%	16%	12%	7%	28%	23%	
Total %	2%	2%	4%	3%	2%	7%	6%	27%
3	5	3	6	17	14	10	11	66
Column %	16%	20%	21%	29%	39%	21%	30%	
Row %	8%	5%	9%	26%	21%	15%	17%	
Total %	2%	1%	2%	7%	6%	4%	4%	26%
4	6	2	5	15	15	15	10	68
Column %	19%	13%	18%	26%	42%	31%	27%	
Row %	9%	3%	7%	22%	22%	22%	15%	
Total %	2%	1%	2%	6%	6%	6%	4%	27%
Totals	32	15	28	58	36	48	37	254
Total %	13%	6%	11%	23%	14%	19%	15%	100%

## Lästig – angenehm

### Omnibus

Statistic	Statistics: PROTOTYP(4) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	44,30639	df=18	p=,00052
M-L Chi-square	44,27657	df=18	p=,00053

### Posthoc-Tests inklusive Bonferroni-Holms-Korrektur Prototyp 1-2

Statistic	Statistics: PROTOTYP(2) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	5,673170	df=6	p=,46078
M-L Chi-square	6,122444	df=6	p=,40961

### Prototyp 1-3

Statistic	Statistics: PROTOTYP(2) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	8,721977	df=6	p=,18983
M-L Chi-square	9,147865	df=6	p=,16543

### Prototyp 1-4

Statistic	Statistics: PROTOTYP(2) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	29,01667	df=6	p=,00006 (0,0003)
M-L Chi-square	31,85316	df=6	p=,00002 (0,0001)

### Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	3,579951	df=6	p=,73330
M-L Chi-square	3,652710	df=6	p=,72355

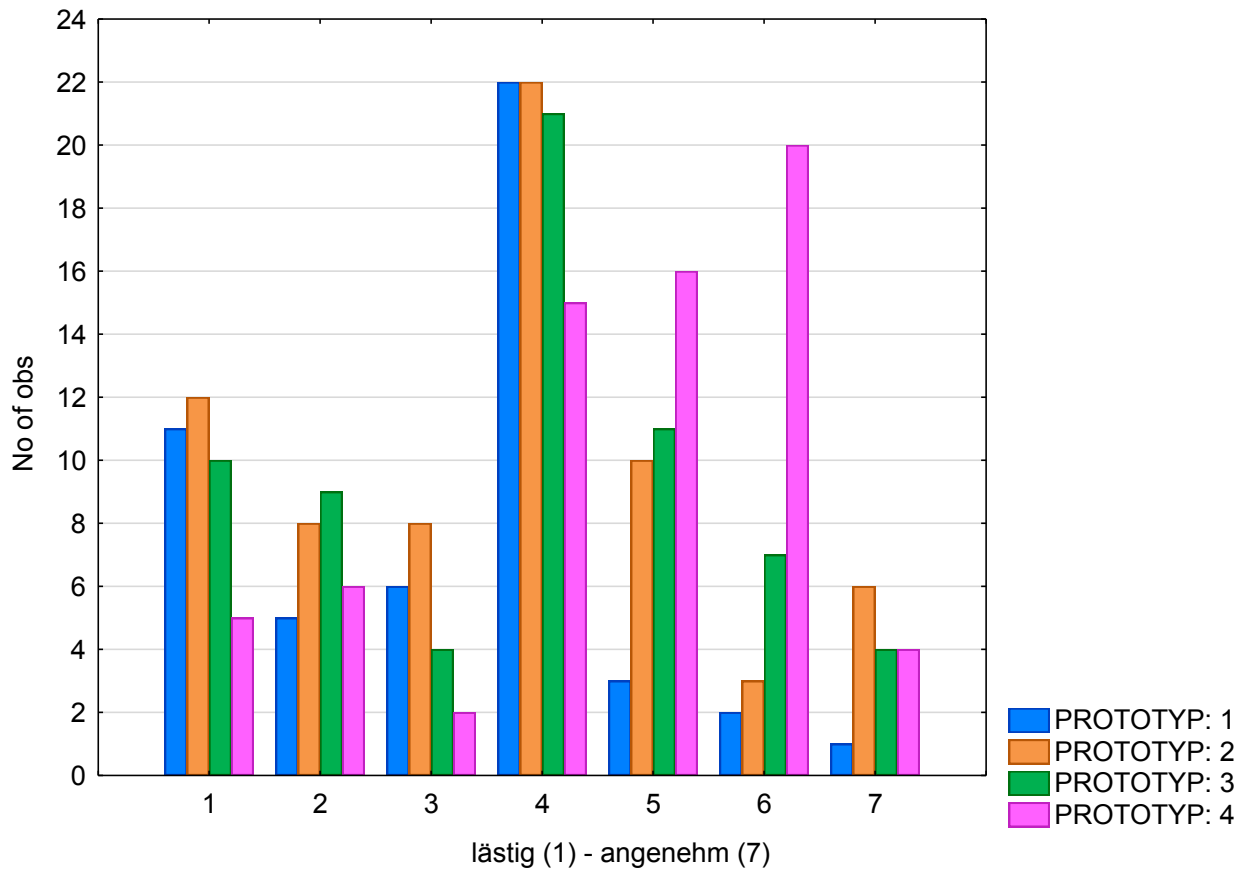
### Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x lästig (1) - angenehm (7)(		
	Chi-square	df	p
Pearson Chi-square	22,43612	df=6	p=,00101 (0,00505)
M-L Chi-square	24,30940	df=6	p=,00046 (0,0023)

### Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x lästig (1) -		
	Chi-square	df	p
Pearson Chi-square	11,09114	df=6	p=,08560
M-L Chi-square	11,41556	df=6	p=,07635

Histogram of lästig (1) - angenehm (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_SBB)								
Marked cells have counts > 10								
PROTOTYP	lästig (1) - angenehm (7) 1	lästig (1) - angenehm (7) 2	lästig (1) - angenehm (7) 3	lästig (1) - angenehm (7) 4	lästig (1) - angenehm (7) 5	lästig (1) - angenehm (7) 6	lästig (1) - angenehm (7) 7	Row Totals
1	11	5	6	22	3	2	1	50
Column %	29%	18%	30%	28%	8%	6%	7%	
Row %	22%	10%	12%	44%	6%	4%	2%	
Total %	4%	2%	2%	9%	1%	1%	0%	20%
2	12	8	8	22	10	3	6	69
Column %	32%	29%	40%	28%	25%	9%	40%	
Row %	17%	12%	12%	32%	14%	4%	9%	
Total %	5%	3%	3%	9%	4%	1%	2%	27%
3	10	9	4	21	11	7	4	66
Column %	26%	32%	20%	26%	28%	22%	27%	
Row %	15%	14%	6%	32%	17%	11%	6%	
Total %	4%	4%	2%	8%	4%	3%	2%	26%
4	5	6	2	15	16	20	4	68
Column %	13%	21%	10%	19%	40%	63%	27%	
Row %	7%	9%	3%	22%	24%	29%	6%	
Total %	2%	2%	1%	6%	6%	8%	2%	27%
Totals	38	28	20	80	40	32	15	253
Total %	15%	11%	8%	32%	16%	13%	6%	100%

## Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen

2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in A)						
Marked cells have counts > 10						
Exclude condition: v6 =1						
PROTOTYP	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen 1	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen 2	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen 3	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen 4	Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen 5	Row Totals
2	13	18	19	17	2	69
Column %	35%	28%	38%	37%	33%	
Row %	19%	26%	28%	25%	3%	
Total %	6%	9%	9%	8%	1%	34%
3	16	25	14	9	2	66
Column %	43%	39%	28%	20%	33%	
Row %	24%	38%	21%	14%	3%	
Total %	8%	12%	7%	4%	1%	33%
4	8	21	17	20	2	68
Column %	22%	33%	34%	43%	33%	
Row %	12%	31%	25%	29%	3%	
Total %	4%	10%	8%	10%	1%	33%
Totals	37	64	50	46	6	203
Total %	18%	32%	25%	23%	3%	100%

Statistics: PROTOTYP(3) x Die Anzeig			
Exclude condition: v6 =1			
Statistic	Chi-square	df	p
Pearson Chi-square	8,798021	df=8	p=,35962
M-L Chi-square	9,171671	df=8	p=,32802

## Hypothese 4: Aussagekraft der Ladeanzeige

- Unberechenbar – Voraussagbar
- Unverständlich – verständlich

### Unberechenbar – vorhersehbar

#### Omnibus

Statistic	Statistics: PROTOTYP(4) x unbereche		
	Chi-square	df	p
Pearson Chi-square	56,45879	df=18	p=,00001
M-L Chi-square	58,09325	df=18	p=,00000

#### Posthoc-Tests inklusive Bonferroni-Holms-Korrektur Prototyp 1-2

Statistic	Statistics: PROTOTYP(2) x unbereche		
	Chi-square	df	p
Pearson Chi-square	36,18885	df=6	p=,00000
M-L Chi-square	40,29179	df=6	p=,00000

#### Prototyp 1-3

Statistic	Statistics: PROTOTYP(2) x unberechenbar (1) - vor		
	Chi-square	df	p
Pearson Chi-square	17,09956	df=6	p=,00892 (0,03568)
M-L Chi-square	17,80661	df=6	p=,00673 (0,02692)

#### Prototyp 1-4

Statistic	Statistics: PROTOTYP(2) x unbereche		
	Chi-square	df	p
Pearson Chi-square	6,396590	df=6	p=,38026
M-L Chi-square	6,552196	df=6	p=,36425

#### Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x unberechenbar (		
	Chi-square	df	p
Pearson Chi-square	15,18152	df=6	p=,01889 (0,05667)
M-L Chi-square	18,36976	df=6	p=,00537 (0,01611)

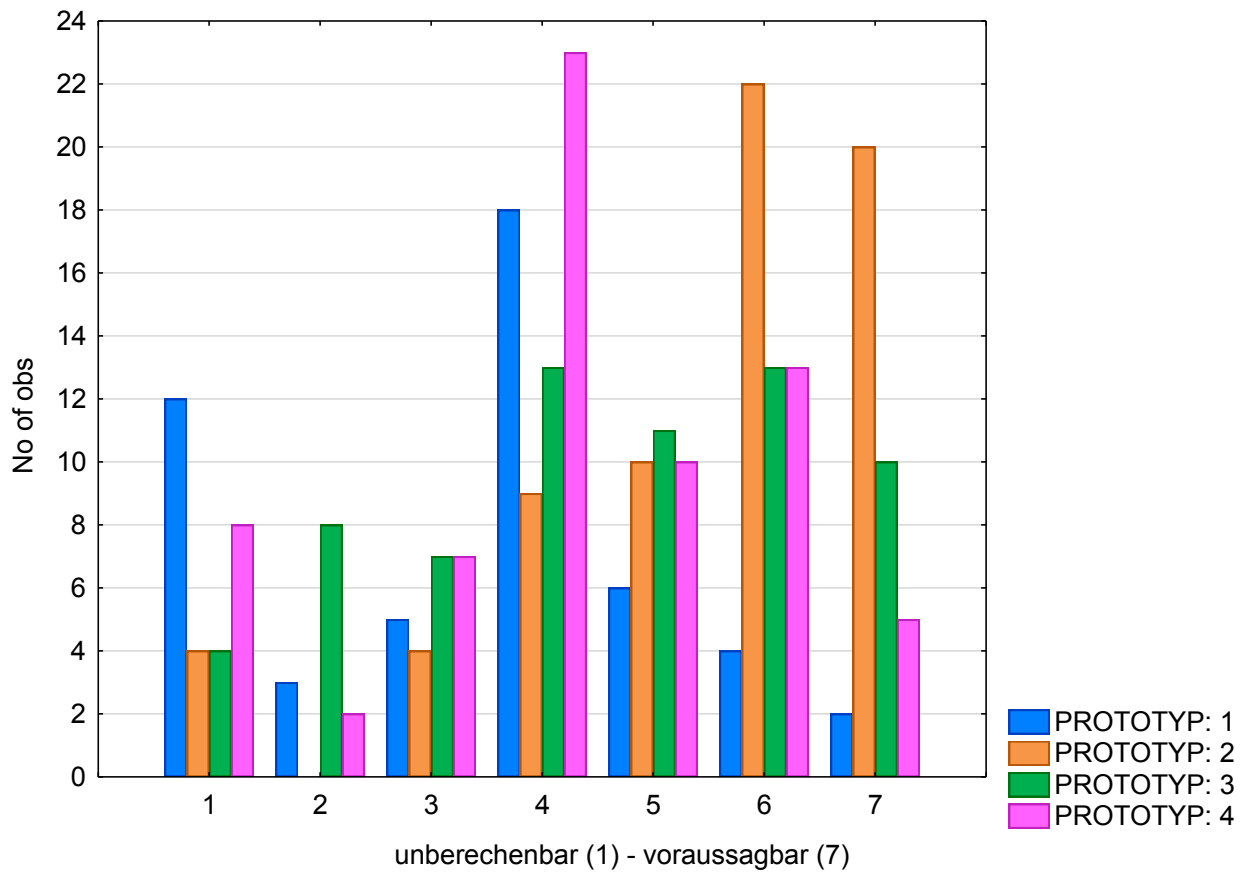
#### Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x unberechenbar (1) - vor		
	Chi-square	df	p
Pearson Chi-square	21,58465	df=6	p=,00144 (0,0072)
M-L Chi-square	23,26795	df=6	p=,00071 (0,00355)

#### Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x unbereche		
	Chi-square	df	p
Pearson Chi-square	9,397639	df=6	p=,15242
M-L Chi-square	9,745514	df=6	p=,13579

Histogram of unberechenbar (1) - voraussagbar (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung\_SBB\_Ladeanzeige)  
 Marked cells have counts > 10

PROTOTYP	unberechenbar (1) - voraussagbar (7) 1	unberechenbar (1) - voraussagbar (7) 2	unberechenbar (1) - voraussagbar (7) 3	unberechenbar (1) - voraussagbar (7) 4	unberechenbar (1) - voraussagbar (7) 5	unberechenbar (1) - voraussagbar (7) 6	unberechenbar (1) - voraussagbar (7) 7	Row Totals
1	12	3	5	18	6	4	2	50
Column %	43%	23%	22%	29%	16%	8%	5%	
Row %	24%	6%	10%	36%	12%	8%	4%	
Total %	5%	1%	2%	7%	2%	2%	1%	20%
2	4	0	4	9	10	22	20	69
Column %	14%	0%	17%	14%	27%	42%	54%	
Row %	6%	0%	6%	13%	14%	32%	29%	
Total %	2%	0%	2%	4%	4%	9%	8%	27%
3	4	8	7	13	11	13	10	66
Column %	14%	62%	30%	21%	30%	25%	27%	
Row %	6%	12%	11%	20%	17%	20%	15%	
Total %	2%	3%	3%	5%	4%	5%	4%	26%
4	8	2	7	23	10	13	5	68
Column %	29%	15%	30%	37%	27%	25%	14%	
Row %	12%	3%	10%	34%	15%	19%	7%	
Total %	3%	1%	3%	9%	4%	5%	2%	27%
Totals	28	13	23	63	37	52	37	253
Total %	11%	5%	9%	25%	15%	21%	15%	100%

## Unverständlich – verständlich

### Omnibus

Statistic	Statistics: PROTOTYP(4) x unverständl		
	Chi-square	df	p
Pearson Chi-square	106,9890	df=18	p=,00000
M-L Chi-square	108,6366	df=18	p=,00000

### Posthoc-Tests inklusive Bonferroni-Holms-Korrektur Prototyp 1-2

Statistic	Statistics: PROTOTYP(2) x unverständl		
	Chi-square	df	p
Pearson Chi-square	52,27077	df=6	p=,00000
M-L Chi-square	60,23852	df=6	p=,00000

### Prototyp 1-3

Statistic	Statistics: PROTOTYP(2) x unverständl		
	Chi-square	df	p
Pearson Chi-square	51,81921	df=6	p=,00000
M-L Chi-square	62,70098	df=6	p=,00000

### Prototyp 1-4

Statistic	Statistics: PROTOTYP(2) x unverständl		
	Chi-square	df	p
Pearson Chi-square	50,49205	df=6	p=,00000
M-L Chi-square	57,00703	df=6	p=,00000

### Prototyp 2-3

Statistic	Statistics: PROTOTYP(2) x unverständlich (1)		
	Chi-square	df	p
Pearson Chi-square	16,31593	df=6	p=,01216 (0,0365)
M-L Chi-square	18,81829	df=6	p=,00448 (0,0134)

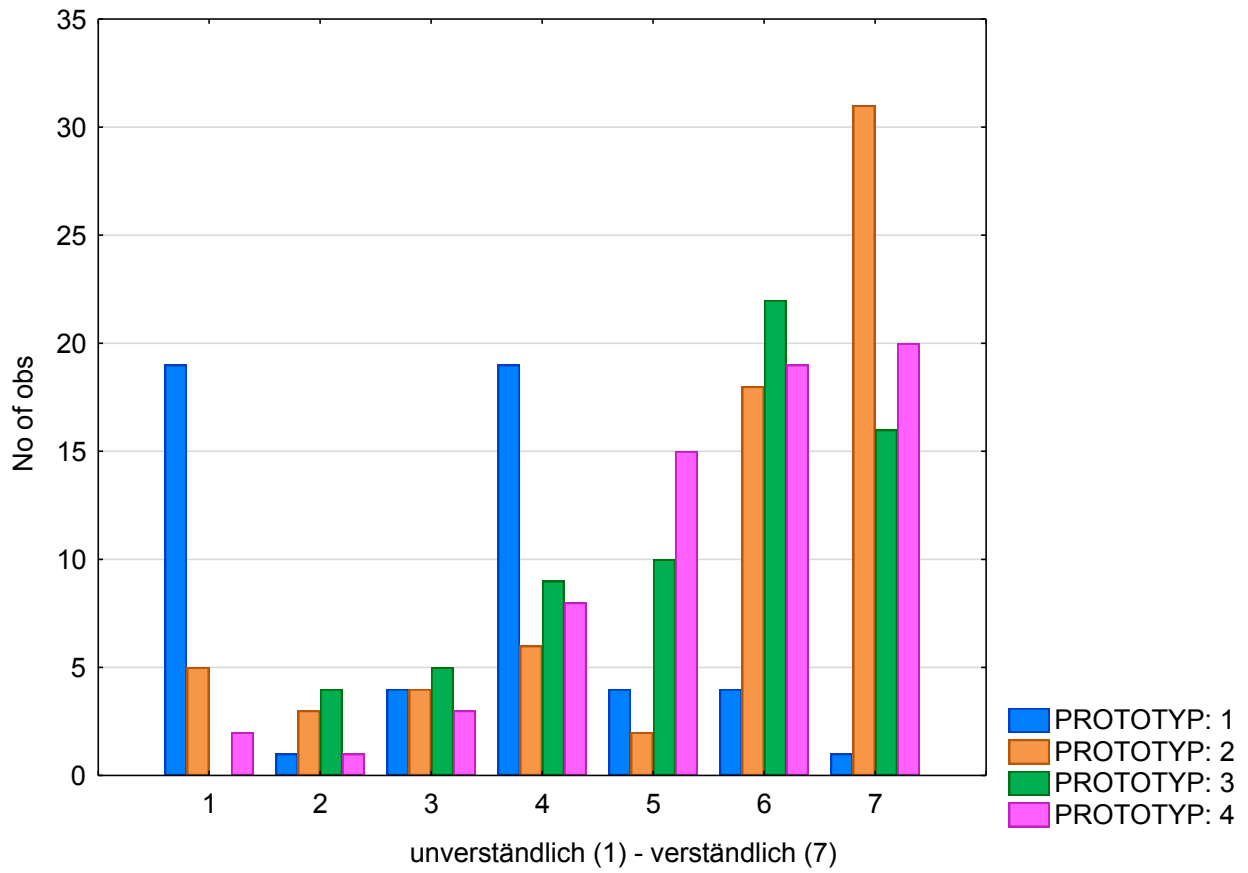
### Prototyp 2-4

Statistic	Statistics: PROTOTYP(2) x unverständlich (1) - v		
	Chi-square	df	p
Pearson Chi-square	15,04854	df=6	p=,01988 (0,03976)
M-L Chi-square	16,46770	df=6	p=,01145 (0,0229)

### Prototyp 3-4

Statistic	Statistics: PROTOTYP(2) x unverständl		
	Chi-square	df	p
Pearson Chi-square	5,994265	df=6	p=,42383
M-L Chi-square	6,906233	df=6	p=,32961

Histogram of unverständlich (1) - verständlich (7); categorized by PROTOTYP



2-Way Summary Table: Observed Frequencies (Datensatz: 3 Fälle ausgeschlossen auf Grund Frage 2 in Auswertung_SBB_Ladeanzeige)								
Marked cells have counts > 10								
PROTOTYP	unverständlich (1) - verständlich (7) 1	unverständlich (1) - verständlich (7) 2	unverständlich (1) - verständlich (7) 3	unverständlich (1) - verständlich (7) 4	unverständlich (1) - verständlich (7) 5	unverständlich (1) - verständlich (7) 6	unverständlich (1) - verständlich (7) 7	Row Totals
1	19	1	4	19	4	4	1	52
Column %	73%	11%	25%	45%	13%	6%	1%	
Row %	37%	2%	8%	37%	8%	8%	2%	
Total %	7%	0%	2%	7%	2%	2%	0%	20%
2	5	3	4	6	2	18	31	69
Column %	19%	33%	25%	14%	6%	29%	46%	
Row %	7%	4%	6%	9%	3%	26%	45%	
Total %	2%	1%	2%	2%	1%	7%	12%	27%
3	0	4	5	9	10	22	16	66
Column %	0%	44%	31%	21%	32%	35%	24%	
Row %	0%	6%	8%	14%	15%	33%	24%	
Total %	0%	2%	2%	4%	4%	9%	6%	26%
4	2	1	3	8	15	19	20	68
Column %	8%	11%	19%	19%	48%	30%	29%	
Row %	3%	1%	4%	12%	22%	28%	29%	
Total %	1%	0%	1%	3%	6%	7%	8%	27%
Totals	26	9	16	42	31	63	68	255
Total %	10%	4%	6%	16%	12%	25%	27%	100%

# G\*Power Berechnung

## Hypothese 2

Variable1	Prototyp -1	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	ATTRAKTIV	ATTRAKTIV	ATTRAKTIV	ATTRAKTIV
Pearson Chi <sup>2</sup>	42.54	8.13	26.74	23.81
ML Chi <sup>2</sup>	42.78	8.29	28.6	25.6
n	203	135	137	134
df	12	6	6	6
alpha-level	0.05	0.05	0.05	0.05
power Pearson	0.998381606	0.53852553	0.986413008	0.973560995
power ML	0.998467717	0.548126311	0.991233477	0.982330523

Variable1	Prototyp -1	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	KONVENTIONELL	KONVENTIONELL	KONVENTIONELL	KONVENTIONELL
Pearson Chi <sup>2</sup>	100.17	8.81	70.56	58.16
ML Chi <sup>2</sup>	106.99	8.98	83	65.72
n	203	135	137	134
df	12	6	6	6
alpha-level	0.05	0.05	0.05	0.05
power Pearson	1	0.57860845	0.999999942	0.99999733
power ML	1	0.588323822	0.999999999	0.999999735

Variable1	Prototyp -1	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	SYMPATHISCH	SYMPATHISCH	SYMPATHISCH	SYMPATHISCH
Pearson Chi <sup>2</sup>	45.62	1.84	30.98	29.44
ML Chi <sup>2</sup>	46.68	1.88	33.95	32.43
n	203	135	137	134
df	12	6	6	6
alpha-level	0.05	0.05	0.05	0.05
power Pearson	0.999205297	0.137265997	0.995075693	0.992833819
power ML	0.99938072	0.139485575	0.997657507	0.996563216

Variable1	Prototyp -1	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	LANGWEILIG	LANGWEILIG	LANGWEILIG	LANGWEILIG
Pearson Chi^2	51.13	5.38	40.1	24.55
ML Chi^2	53.68	5.8	46.15	26.5
n	203	135	137	134
df	12	6	6	6
alpha-level	0.05	0.05	0.05	0.05
power Pearson	0.999787854	0.361559454	0.999532089	0.977585922
power ML	0.999887066	0.389558441	0.999911564	0.985634896

Variable1	Prototyp -1	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	ATTRAKTIV EMPFUNDEN	ATTRAKTIV EMPFUNDEN	ATTRAKTIV EMPFUNDEN	ATTRAKTIV EMPFUNDEN
Pearson Chi^2	38.54	4.36	33.88	17.84
ML Chi^2	39.54	4.41	35.77	18.58
n	203	135	137	134
df	12	6	6	6
alpha-level	0.05	0.05	0.05	0.05
power Pearson	0.996052672	0.293533033	0.99761546	0.907855206
power ML	0.996830056	0.296847689	0.998531583	0.920364894

### Hypothese 3

Variable1	Prototyp	Prototyp 1-2	Prototyp 1-3	Prototyp 1-4	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	NUTZLOS	NUTZLOS	NUTZLOS	NUTZLOS	NUTZLOS	NUTZLOS	NUTZLOS
Pearson Chi <sup>2</sup>	33.82	18.33	17.06	14.35	3.03	11.55	3.63
ML Chi <sup>2</sup>	35.1	20.36	18.33	14.76	3.05	11.98	3.72
n	255	121	118	120	135	137	134
df	18	6	6	6	6	6	6
alpha-level	0.05	0.05	0.05	0.05	0.05	0.05	0.05
power Pearson	0.975433704	0.916315632	0.892859091	0.823690414	0.207602147	0.71857244	0.245709838
power ML	0.980418882	0.944538391	0.916315632	0.836044873	0.20884929	0.73718205	0.251533957

Variable1	Prototyp	Prototyp 1-2	Prototyp 1-3	Prototyp 1-4	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	VERWIRREND	VERWIRREND	VERWIRREND	VERWIRREND	VERWIRREND	VERWIRREND	VERWIRREND
Pearson Chi <sup>2</sup>	58.13	34.72	27.12	30.73	12.87	11.89	1.56
ML Chi <sup>2</sup>	64.05	41.26	32.38	35.9	13.18	12.24	1.57
n	254	120	117	119	135	137	134
df	18	6	6	6	6	6	6
alpha-level	0.05	0.05	0.05	0.05	0.05	0.05	0.05
power Pearson	0.999827933	0.998075491	0.987564765	0.994763908	0.772799706	0.733363153	0.122062002
power ML	0.999956623	0.999658048	0.996519982	0.998580219	0.784304063	0.747989033	0.122594645

Variable1	Prototyp	Prototyp 1-2	Prototyp 1-3	Prototyp 1-4	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	LÄSTIG	LÄSTIG	LÄSTIG	LÄSTIG	LÄSTIG	LÄSTIG	LÄSTIG
Pearson Chi^2	44.31	5.67	8.72	29.02	3.58	22.44	11.09
ML Chi^2	44.28	6.12	9.15	31.85	3.65	24.31	11.42
n	253	119	116	118	135	137	134
df	18	6	6	6	6	6	6
alpha-level	0.05	0.05	0.05	0.05	0.05	0.05	0.05
power Pearson	0.996619841	0.38090732	0.573414125	0.992071729	0.242484941	0.96430731	0.697642855
power ML	0.996599334	0.410766701	0.597911199	0.996028605	0.247001978	0.976347232	0.712764754

### Hypothese 4

Variable1	Prototyp	Prototyp 1-2	Prototyp 1-3	Prototyp 1-4	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	UNVERSTÄNDLICH	UNVERSTÄNDLICH	UNVERSTÄNDLICH	UNVERSTÄNDLICH	UNVERSTÄNDLICH	UNVERSTÄNDLICH	UNVERSTÄNDLICH
Pearson Chi <sup>2</sup>	106.99	52.27	51.82	50.49	16.32	15.05	5.99
ML Chi <sup>2</sup>	108.64	60.24	62.7	57	18.82	16.47	6.91
n	255	121	118	120	135	137	134
df	18	6	6	6	6	6	6
alpha-level	0.05	0.05	0.05	0.05	0.05	0.05	0.05
power	0.999999999	0.999984704	0.99998256	0.999974351	0.876747794	0.844351018	0.402167789
Pearson							
power ML	1	0.999998576	0.999999328	0.999996219	0.924089581	0.880169881	0.462328101

Variable1	Prototyp	Prototyp 1-2	Prototyp 1-3	Prototyp 1-4	Prototyp 2-3	Prototyp 2-4	Prototyp 3-4
Variable2	UNBERECHENBAR	UNBERECHENBAR	UNBERECHENBAR	UNBERECHENBAR	UNBERECHENBAR	UNBERECHENBAR	UNBERECHENBAR
Pearson Chi <sup>2</sup>	56.46	36.19	17.1	6.4	15.18	21.58	9.4
ML Chi <sup>2</sup>	58.09	40.29	17.81	6.55	18.37	23.27	9.75
n	253	119	116	118	135	137	134
df	18	6	6	6	6	6	6
alpha-level	0.05	0.05	0.05	0.05	0.05	0.05	0.05
power	0.99974877	0.99868319	0.89367618	0.429191522	0.847960898	0.957076247	0.611772556
Pearson							
power ML	0.999826356	0.999555429	0.907313912	0.438999473	0.916975367	0.970214505	0.630691153

## Umfrage über Wartezeiten und Ladevorgang bei SBB Ticketshop

### Fast geschafft, jetzt noch ein paar Fragen ...

Hast du eine Anzeige dieses Ladevorgangs gesehen?  
 Falls ja, welche? (siehe Abbildung unten) \*

- A  
 B  
 C  
 Nein, ich habe keine Abbildungen gesehen

<p>Hinfahrt <b>A</b></p> <p>Zürich HB Bern</p> <p>2. Klasse</p> <p>Gültig: Fr., 9. Sep. 2016</p> <p>Billette ab CHF 25.00</p> 	<p>Hinfahrt <b>B</b></p> <p>Zürich HB Bern</p> <p>2. Klasse</p> <p>Gültig: Fr., 9. Sep. 2016</p> <p>Billette ab CHF 25.00</p> 	<p>Hinfahrt <b>C</b></p> <p>Zürich HB Bern</p> <p>2. Klasse</p> <p>Gültig: Fr., 9. Sep. 2016</p> <p>Billette ab CHF 25.00</p> 
<p>Gesamtpreis CHF --.--</p> <p>Zur-Kasse-</p>	<p>Gesamtpreis CHF --.--</p> <p>Tarife werden berechnet</p> <p>Zur-Kasse-</p>	<p>Gesamtpreis CHF --.--</p> <p>Zur-Kasse-</p>

Wieviele Sekunden hat deiner Einschätzung nach die Ladezeit gedauert? \*

Wie empfandst du die Dauer des Ladevorgangs? \*

- |                                       |                       |                       |                       |                       |                       |
|---------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                                       | 1                     | 2                     | 3                     | 4                     | 5                     |
|                                       | trifft gar nicht zu   | trifft eher nicht zu  | teils-teils           | trifft eher zu        | trifft absolut zu     |
| Ich empfand die Dauer als angemessen. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

**Wie hast du die Anzeige des Ladevorgangs empfunden? \***

	1	2	3	4	5	6	7	
langsam	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	schnell
nutzlos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	nützlich
verwirrend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	übersichtlich
unattraktiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	attraktiv
unberechenbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	voraussagbar
konventionell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	originell
lästig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	angenehm
unsympathisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sympathisch
langweilig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	spannend
unverständlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	verständlich

**Im Vergleich mit deinen sonstigen Erfahrungen im Internet, wie bewertest du den Ladevorgang auf der Testseite? \***

	1 trifft gar nicht zu	2 trifft eher nicht zu	3 teils-teils	4 trifft eher zu	5 trifft absolut zu
Die Anzeige des Ladevorgangs half meine Wartezeit kürzer zu machen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich konnte nachvollziehen wieso diese Anzeige des Ladevorgangs erschien	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Wie empfandest du die Darstellung des Ladevorgangs? \***

	1 trifft gar nicht zu	2 trifft eher nicht zu	3 teils-teils	4 trifft eher zu	5 trifft absolut zu
Ich empfand die Darstellung als optisch attraktiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Während des Ladevorgangs hat sich der Warenkorb aktualisiert.**

**Hast du diese Veränderung bemerkt? \***

- Ja
- Nein
- Ich bin mir nicht mehr sicher

**Wenn ja, hat sich der Warenkorb so verändert wie du es erwartet hast?**

- Ja
- Nein

**Möchtest du uns noch etwas mitteilen?**

**Wir brauchen noch ein paar Daten über dich**

**Geschlecht \***

- Frau
- Mann

**Alter \***

- 25 Jahre oder jünger
- 26-40
- 41-55
- 56-70
- 71 und älter

### Wie bewertest du deine Affinität für Technologie? \*

	1 trifft gar nicht zu	2 trifft eher nicht zu	3 teils-teils	4 trifft eher zu	5 trifft absolut zu
Ich komme im Alltag bei der Bedienung von Computern gut zurecht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Internet ist meine wichtigste Informationsquelle im Alltag	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### Wie oft fährst du mit der SBB? \*

- (fast) täglich
- mehrmals pro Woche
- mehrmals pro Monat
- mehrmals pro Jahr
- (fast) nie

### Wie kaufst du deine Bahnbillette am häufigsten? \*

- Via SBB-Webshop am Computer
- Via SBB-App im Smartphone
- Via Billett-Automat am Bahnhof
- Direkt am Schalter
- Sonstige

### Verlosungsteilnahme

Wir verlosen 100 CHF in Reka-Checks und eine SBB Tageskarte unter den Teilnehmern. Möchtest du teilnehmen? Dann fülle bitte Name und E-Mail-Adresse aus, damit wir dich kontaktieren können, falls du gewinnst. Vielen Dank!

### Möchtest du an der Verlosung teilnehmen?

- ja
- nein

**Wenn ja, gib bitte deine Daten ein. Danke!**

Diese Information wird nicht für andere Zwecke verwendet und nach Benachrichtigung der Gewinner wieder gelöscht

Deine Name

Deine E-Mail-Adresse

» **Umleitung auf Schlussseite von Umfrage Online** ([ändern](#))

## Szenario für die Umfrage

Sie haben heute Göttag und möchten mit Ihrem Götibub und dem Grossvater Zürich nach Bern fahren, um den Bärengraben zu besuchen.

Sie gehen zur SBB-Webseite und suchen die Verbindung die Ihnen am besten passt.

(Die Eingabefelder sind bereits ausgefüllt)

Dann kommen Sie auf die Reise-Optionen-Seite und fügen in den Warenkorb Tickets für alle Reisenden hinzu.

- 2 Erwachsene mit Halbtax
- 1 Kind mit Junior-Karte

Bitte füllen Sie bei allen Reisenden alle Felder aus.

Bitte achten Sie auf die automatische Preis Veränderung im Warenkorb.

Am Ende merken Sie sich bitte den endgültigen Preis.

Gehen Sie zum nächsten Prototypen.

# Pre-Interview

Alter

Geschlecht

Online-Affinität

Reiseverhalten - wie oft wird Zug gefahren

Jeden Tag

Einmal pro Woche

Mehrmals pro Monat

Ab und zu

Sehr selten

Welches oder welche der folgenden Abonnemente des öffentlichen Verkehrs besitzen Sie persönlich?

Ich besitze kein Abonnement

Generalabonnement (inkl. FVP-GA)

Halbtaxabonnement

Verbundabonnement (z.B. ZVV-Netzpass, U-Abo, Libero, Ostwind, A-Welle)

Streckenabonnement

Gleis 7

Anderes Abo wie Juniorkarte, Enkelkarte

Wieso haben Sie dieses Abonnement gewählt?

Wie kaufen Sie Ihr Ticket am häufigsten?

Desktop

App

Billette-Automat

Schalter

Wieso wählen Sie diese Kauf-Methode?

# Post-Interview

Welcher Ladevorgang haben Sie am langsamsten empfunden? Setzen Sie eine Reihenfolge und begründen Sie Ihren Entscheid.

Welcher Ladevorgang haben Sie am schnellsten empfunden? Setzen Sie eine Reihenfolge und begründen Sie Ihren Entscheid.

Welcher Ladevorgang hat Ihnen am Besten gefallen? Setzen Sie eine Reihenfolge und begründen Sie Ihren Entscheid.

Wenn Sie heute eine Entscheidung fällen müssen, welchen Ladevorgang die SBB verwenden soll, was würden Sie auswählen?

---

Fallen Ihnen Beispiele von guten Ladevorgängen ein? Zum Beispiel Swiss?  
/ erklären

Was hat Ihnen an diesen Beispielen gefallen? Begründen Sie Ihren Entscheid.

Weitere Meinungen / Feststellungen

## Usability-Test: Beispiel von Auswertung

ID	Prototyp	Zeit	Screen	Problem	Schweregrad	Ursache	Empfehlung
H1	2	02:50	verbindung	Kontext des Verbindungsscreen ist stark reduziert	2	zu wenig Einbettung	Mehr Einbindung in den Websektekontext
H2	2	03:40	verbindung	Schwierigkeiten in der Bedienung des Prototyps (kein Detailscreen vorhanden)	2		Mehr Einbindung in den Websektekontext
H3	2	06:42	optionen	Mehrmaliger Ladebalken ist schlecht	0	Bug	Bugfix im Prototyp
H4	2	06:50	optionen	Design: Zu grosser Abstand	1	Design	lassen?
H5	2	08:30	optionen	Verwirrende Informationen im Warenkorb	4	Prototypinhalte	Inhalte anpassen
H6	1	16:30	optionen	Verwirrung wegen fehlender Rückmeldung	0	Experimentdesign	
H7	4	22:00	optionen	Reaktion auf Zug-Ladebalken	2	Interessantes Design	Eventuell verwenden
H8	4	22:45	optionen	Züge als Ladebalken nerven	3	Interessantes Design	Eventuell nicht immer verwenden
H9	3	02:45	optionen	Gute Bewertung des detaillierten Feedbacks	3		
H10	3	02:48	optionen	Mehrmaliger gleicher Feedbacktext ist verwirrend	3	Experimentdesign	Darauf achten, dass Feedbacktext treffend ist

# Versuchsplan

## Quasi-experimenteller Test

Der Test hat nur eine einzige Variable: Die vier Prototypen unterscheiden sich bloss was die Visualisierung der Ladeanzeige angeht. Insbesondere ist die Ladedauer und Positionierung bei allen Prototypen genau gleich.

Die visuelle Wahrnehmung der Benutzer haben wir durch die Eliminierung des realen Einkaufsprozesses manipuliert. Der Prozess ist durch ein Storyboard (Bilder) ersetzt. Der Benutzer wird nach dem Lesen des Storyboards direkt auf die Shop-Seite geführt.

In unserem Test berücksichtigen wir nur eine Benutzergruppe. Da die verzögerte Ladedauer auf alle Benutzer zutrifft, haben wir die folgenden Faktoren nicht berücksichtigt: Geschlecht, Herkunft und Bildungsstand, wie oft die Personen mit dem Zug fahren oder ob sie ein General Abonnement besitzen.

## Störfaktoren

Im Weiteren konnten Störfaktoren nicht kontrolliert werden, da wir eine quantitative Online- Umfrage gestartet haben. Mögliche Störfaktoren wie zum Beispiel der Lärmpegel in der unmittelbaren Umgebung des Benutzer, zum Beispiel wenn der Test im Büro ausgefüllt wird, Lichtverhältnisse oder die Geschwindigkeit des Browsers, konnten wir nicht beeinflussen. Die Benutzer wurden via Zufallszahlengenerator einem der vier Prototypen zugewiesen. Wegen der hohen Ausstiegsrate des Prototypen 1, wurden die letzten Benutzer (Reserve) auf den Prototypen 1 zugewiesen.

	Quasi-Experiment
Kausaltheorietische Hypothese vor Versuchsbeginn vorhanden und begründet	+
Experimentelle Variablen manipuliert	+
Alle übrigen Versuchsbedingungen kontrollierbar bzw kontrolliert.	-

## Hypothesen

Wir wollten mit unserer quantitativen Umfrage folgende Hypothesen prüfen:

### **1. Eine Ladeanzeige hilft, dass der Benutzer die Ladedauer als kürzer einschätzt, bzw. wahrnimmt.**

Die Literatur über Responsiveness und User Experience empfehlen immer der User unterhalten und informiert zu halten. Gemäss Miller (1985) möchten die Benutzen Progressindikatoren haben. Der Art, wie dieser Progressindikator sich präsentiert und bewegt, kann die Zeitwahrnehmung der User beeinflussen (Harrison, Amento, Kuznetsov, & Bell, 2007)

### **2. Die Zufriedenheit mit der visuellen Ästhetik der Ladeanzeige ist beim Prototyp 4 am grössten.**

Hintergrund: Benutzer-Zufriedenheit im Bezug auf die Ästhetik der Ladevorgangsanzeige. Prototyp 4 zeigt den Ladevorgang mit einem Zug an, während die anderen Prototypen weniger speziell sind. (Prototyp 1 nicht berücksichtigen)

Gemäss Donald Normans These «Attractive things work better» reagieren Benutzer auf eine visuell attraktive Anzeige besser als auf eine Anzeige die gewöhnlicher scheint. Im Weiteren nimmt der Benutzer drei zentrale Dimensionen wahr wenn er eine Website beurteilt: Inhalt, Usability und Ästhetik (vgl. Thielsch 2008a).

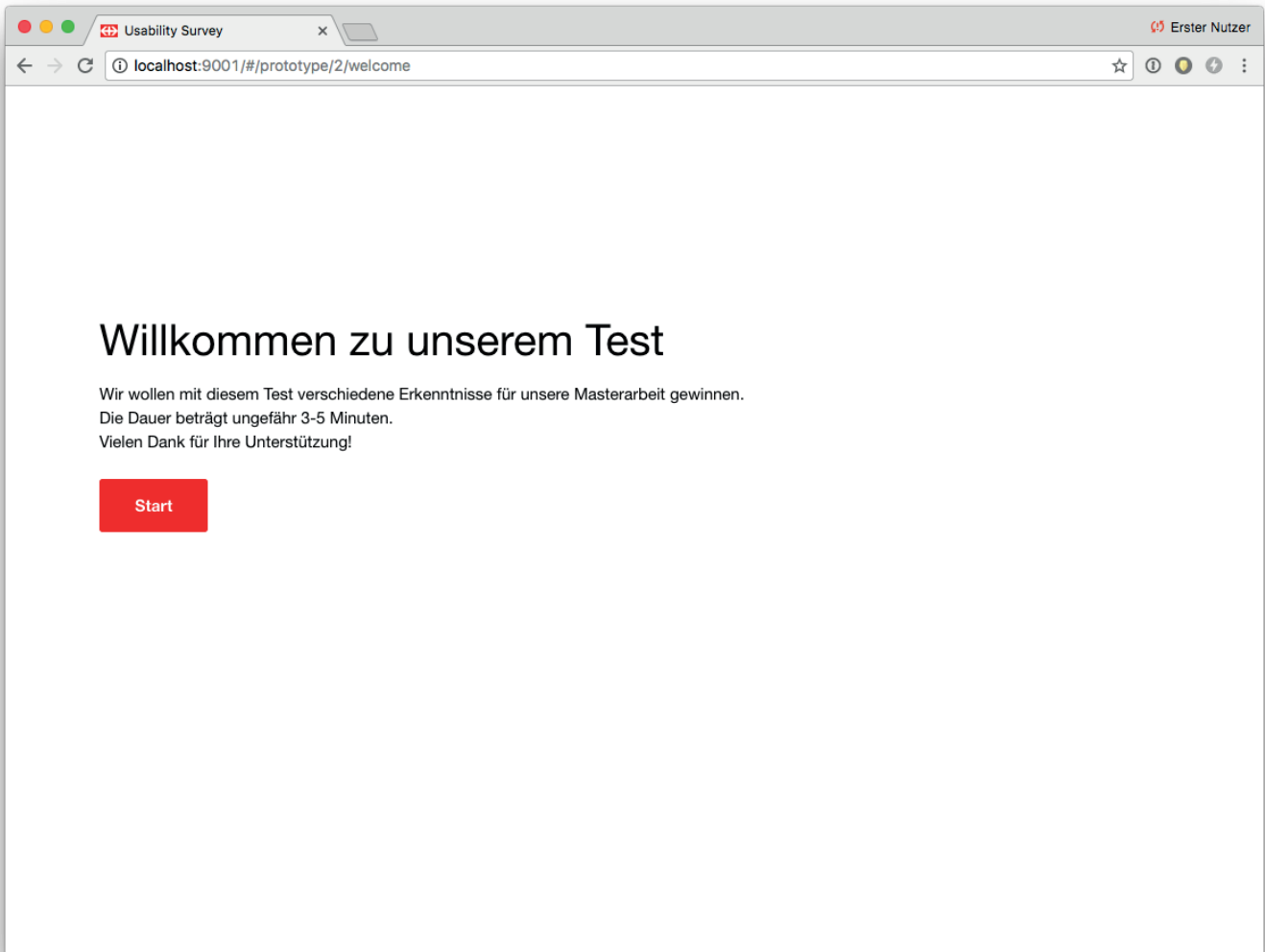
### **3. Die Zufriedenheit der Ladedauer ist beim Prototyp 3 am grössten.**

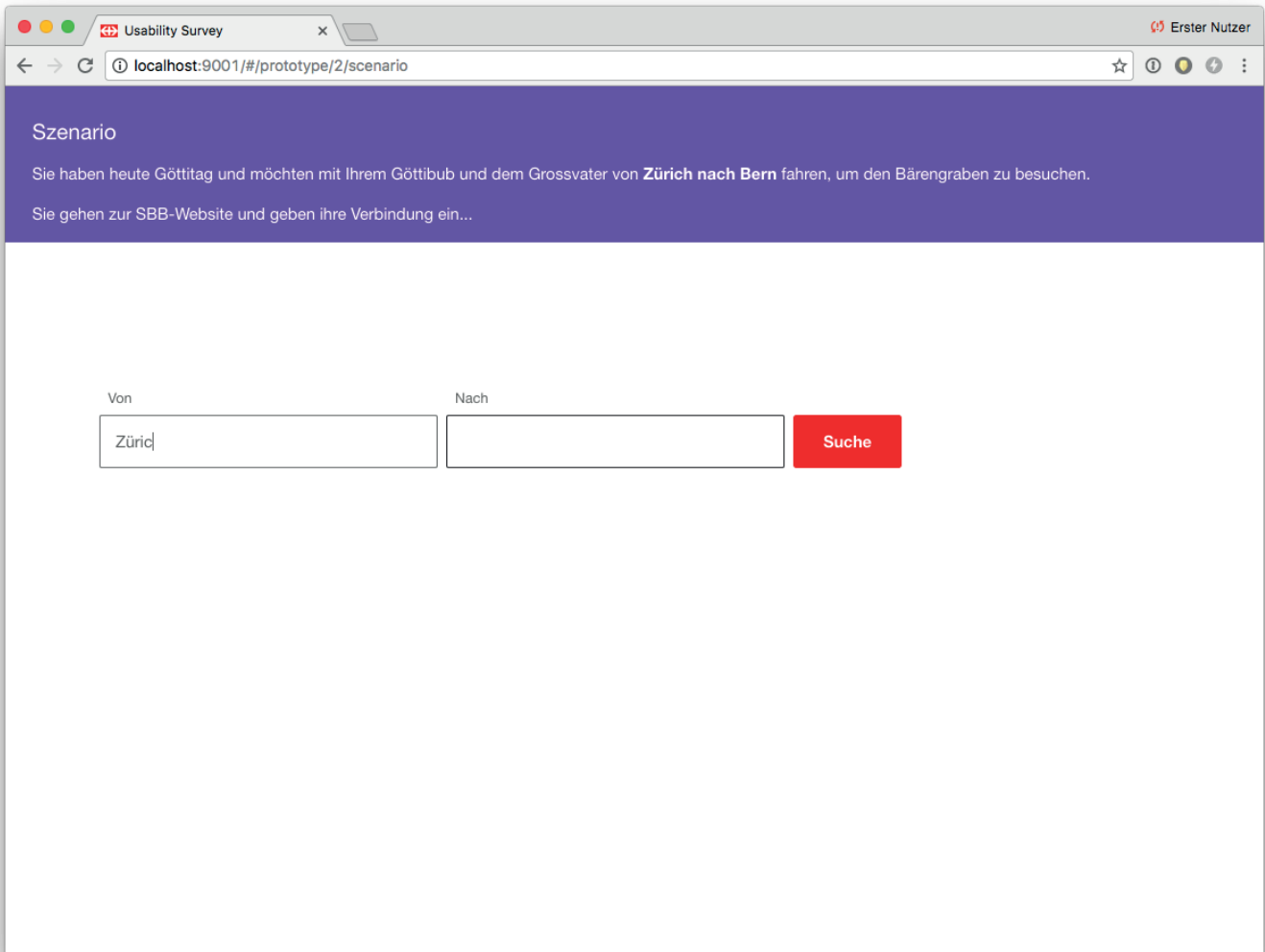
Hintergrund: Benutzer-Zufriedenheit im Bezug auf die Ladedauer. Bei Prototyp 3 erhalten Benutzer Feedback sowohl visuell als auch in der Form von relevanter textuellen Beschreibungen.

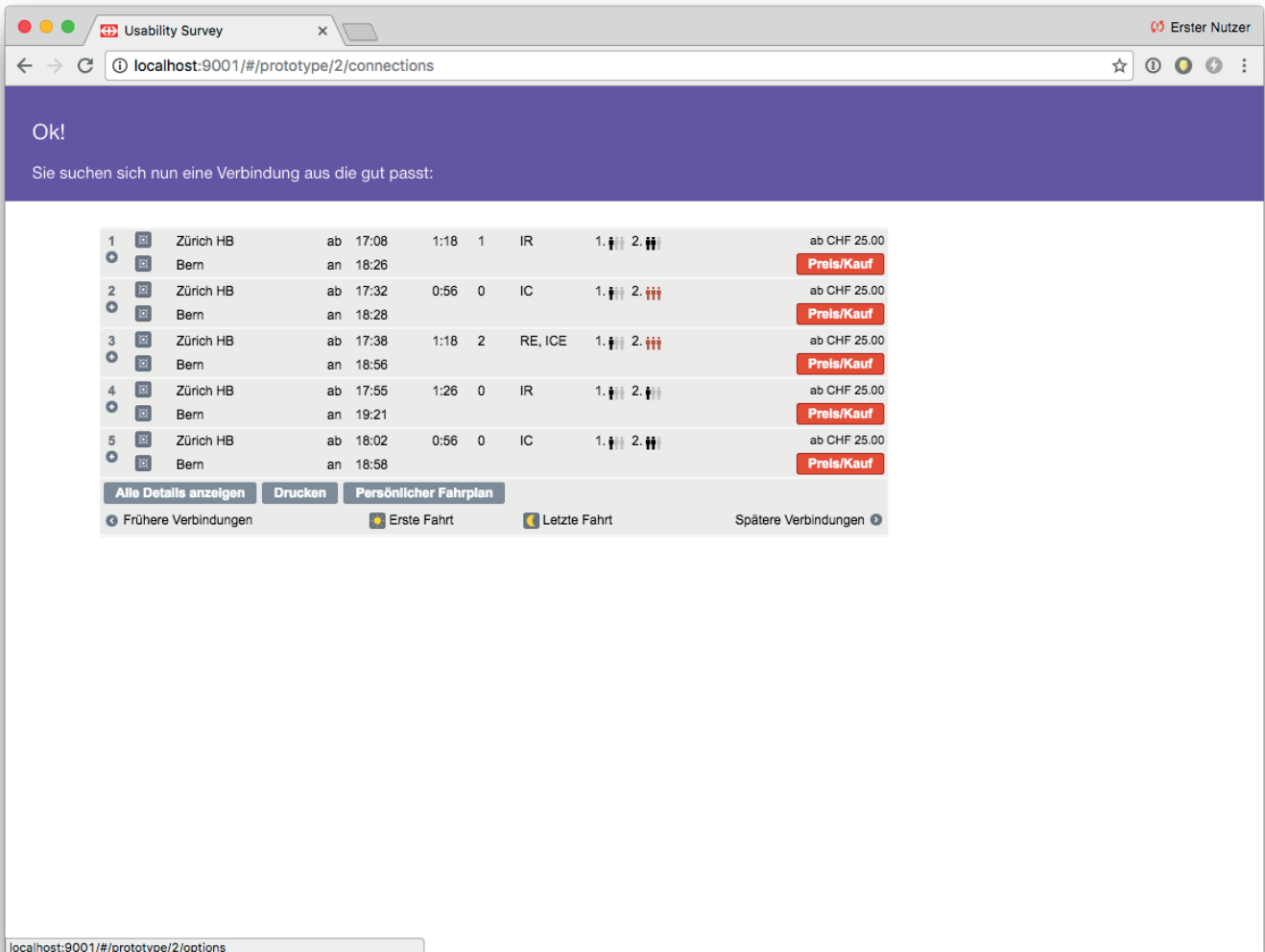
### **4. Die Aussagekraft der Ladeanzeige ist bei Prototyp 3 am stärksten (bei Prototyp 1 am schwächsten).**

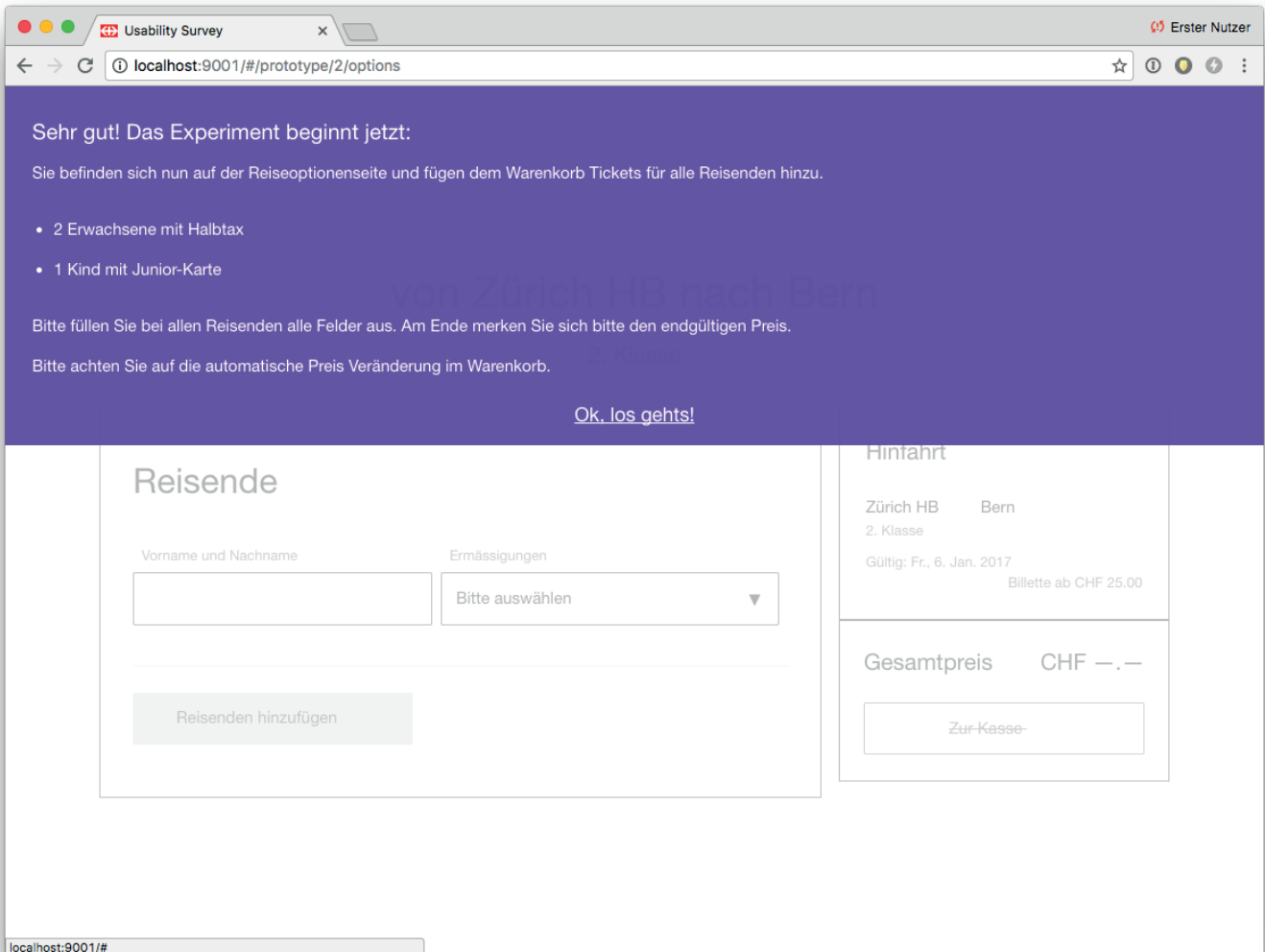
Hintergrund: Der Benutzer braucht eine aussagekräftige Ladeanzeige und die Rückmeldung trägt zum allgemeinen Verständnis der Benutzung der Applikation bei.

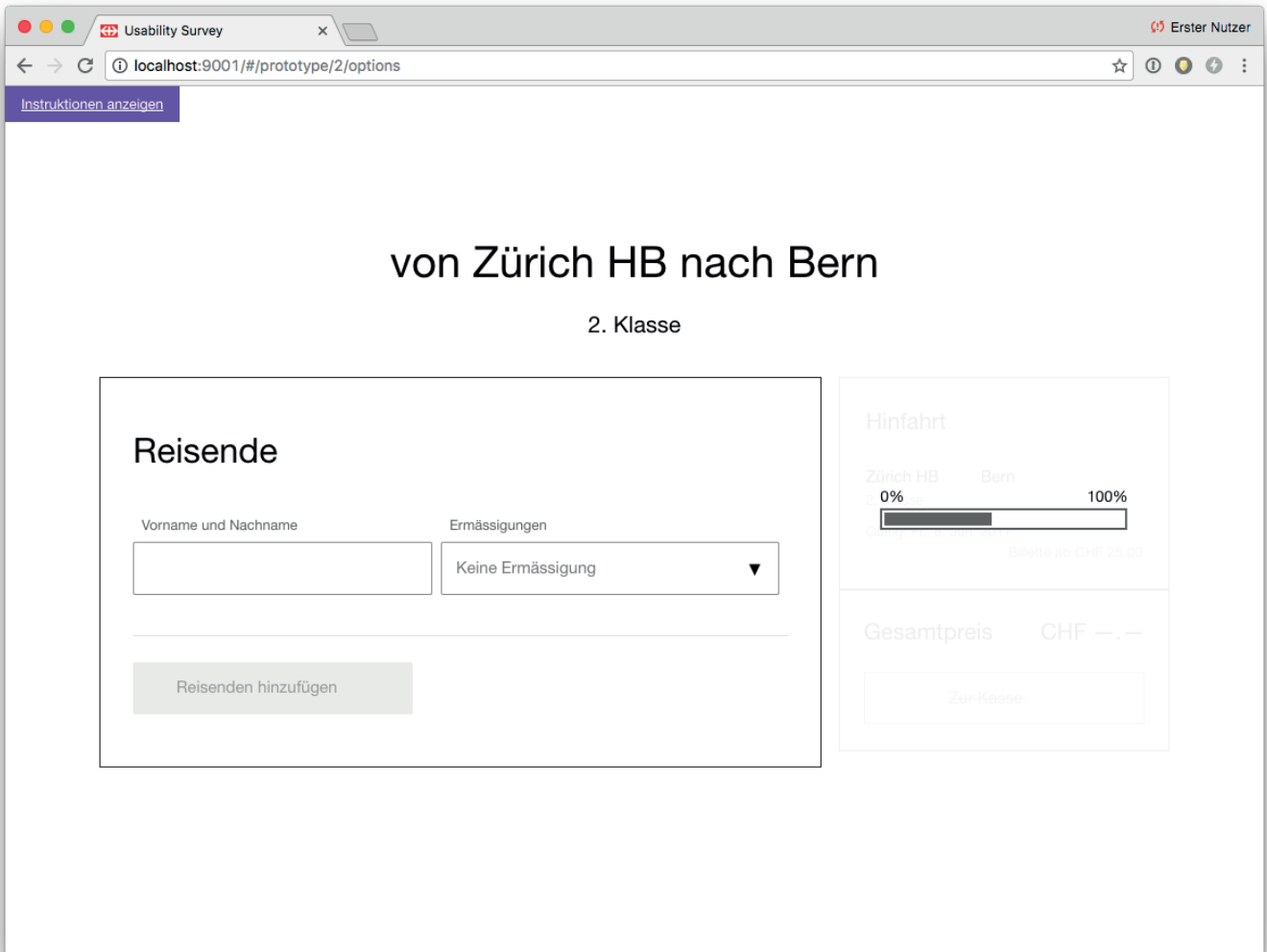
# Prototyp - V1.0

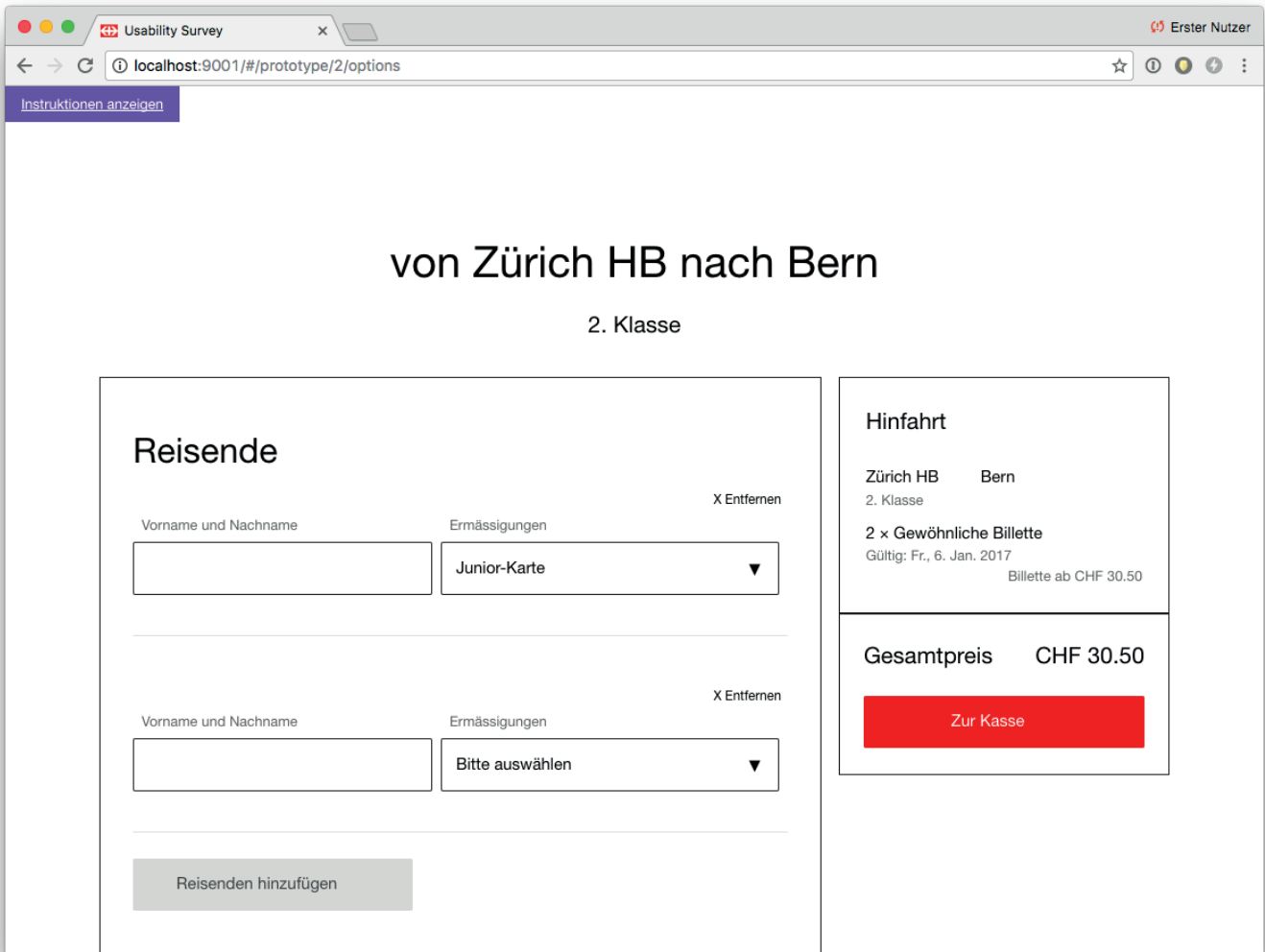




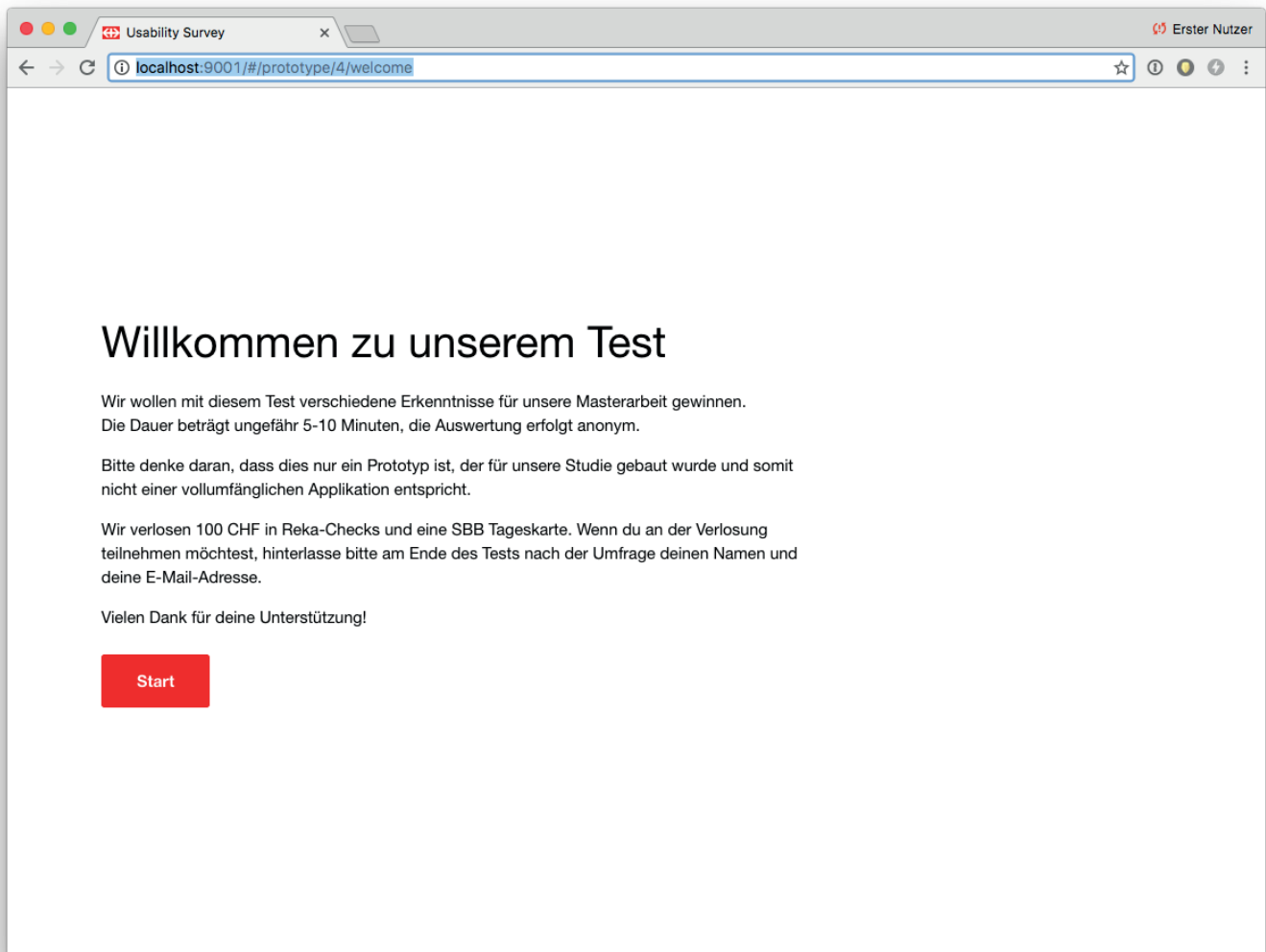


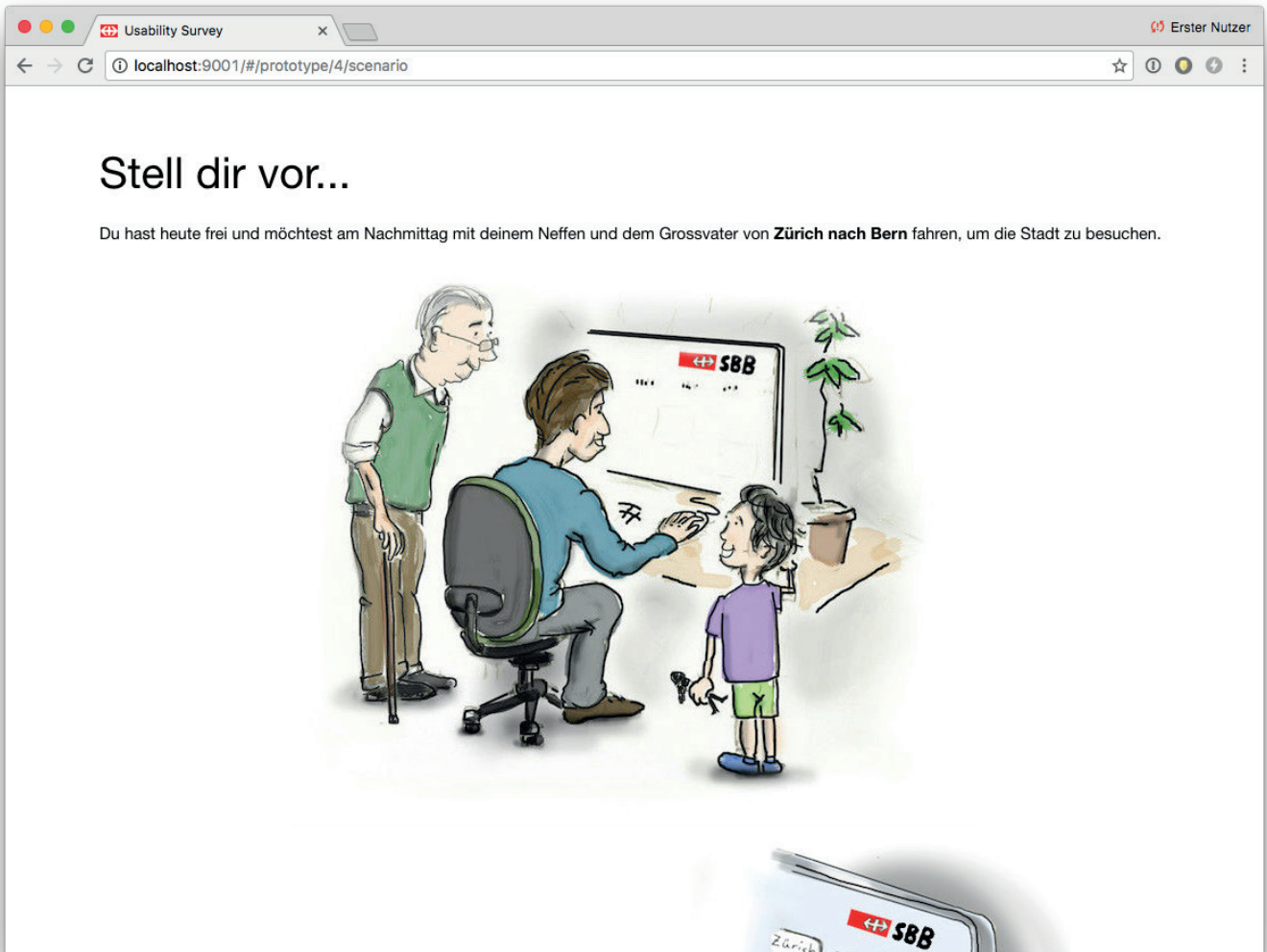






# Prototyp - V2.1

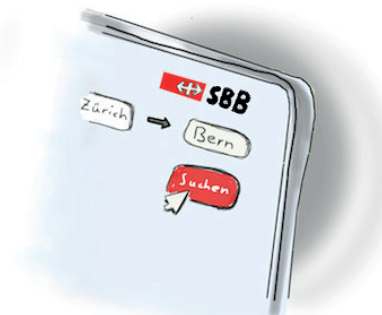





Usability Survey x Erster Nutzer

localhost:9001/#/prototype/4/scenario

Du gehst zur SBB-Website und gibst deine Suchanfrage für die Verbindung ein.



Nun klickst du auf eine der angezeigten Verbindungen um weiter zu kommen.



Weiter ...

