



Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems

Bachelorarbeit

von

Daniel Braune

geboren in
Langenfeld

eingereicht am

Institut für Informatik
Prof. Dr. Martin Mauve
Heinrich-Heine-Universität Düsseldorf

September 2016

Betreuer:
Tobias Krauthoff, M.Sc.

Zusammenfassung

Viele Internetnutzer führen online Diskussionen, was mit Plattformen wie Foren, Blogs oder Social-Media Seiten möglich ist. Problematisch ist jedoch, dass diese Möglichkeiten oft unstrukturiert sind, gerade wenn mehrere hundert Nutzer daran teilnehmen. Aussagen des Nutzers, die er zu einem früheren Zeitpunkt in der Diskussion, oder auch in anderen Diskussionen, geäußert hat, bleiben oftmals unberücksichtigt. Insbesondere die Darstellung der eigenen Meinung mit Freitext ist hierbei oft ungeordnet, wobei Probleme wie Balkanisierung, Spam und Redundanz entstehen. Dieser Problemstellung nimmt sich D-BAS an, ein neues, dialog-basiertes Argumentationssystem, welches es Nutzern ermöglicht, Diskussionen online zu führen, während es diese vereinfacht und übersichtlicher gestaltet. Der Nutzer wählt hierbei während der Diskussion zwischen verschiedenen Antwortmöglichkeiten, die ihm das System vorgibt.

In dieser Arbeit soll herausgefunden werden, ob D-BAS dem Nutzer durch diese Antwortmöglichkeiten die selben Interaktionsmöglichkeiten bietet, die dieser in einer realen Diskussion zwischen Personen ebenfalls hätte. Dazu wurden zwei Studien durchgeführt, in denen das System auf genau diese Interaktionsmöglichkeiten getestet wurde. Das Ergebnis dieser Tests hat gezeigt, dass es während der Studien zu Verständnisproblemen bezüglich der Verknüpfungen der einzelnen Aussagen des Systems gekommen ist. Des Weiteren war es anfänglich für den Nutzer schwer, seine eigene Meinung auf eine der Antwortmöglichkeiten von D-BAS zu übertragen, da dieser sich nicht immer sicher war, welche der Möglichkeiten seine Meinung am passendsten widerspiegelt. Hierbei wurden lange Sätze sowie schwere Formulierungen in den Antwortmöglichkeiten kritisiert. Durch die Einführung des Colour-Codings sowie der Verkürzung der jeweiligen Antwortmöglichkeiten durch die Ersetzung von Argumenten durch kurze, farblich markierte Phrasen, wurden nach der ersten Studie bereits die genannten Verbesserungen am System durchgeführt, welche im zweiten Durchlauf getestet wurden. Nach Abschluss dieses Durchlaufs wurden weitere Verbesserungen vorgeschlagen, welche eine Ergänzung des Colour-Codings sowie die grammatikalische Umformulierung der Verknüpfungen zwischen verschiedenen Aussagen im System beinhalten. Der Fokus wurde darauf gelegt, dass D-BAS in Zukunft leicht verständlich wird und somit für den gewöhnlich, durchschnittlichen Internetnutzer einfach zu bedienen sein soll.

Danksagung

An dieser Stelle möchte ich mich bei allen Menschen bedanken, die mich im Rahmen meiner Bachelorarbeit begleitet und unterstützt haben.

Ganz besonders bedanken möchte ich mich bei Prof. Dr. Mauve sowie Tobias Krauthoff, die mich sowohl fachlich als auch persönlich bei der Anfertigung dieser Arbeit begleitet haben.

Außerdem danke ich meinen Eltern Gaby und Michael Braune, die mir durch ihre Unterstützung das Studium ermöglicht haben.

Ein besonderer Dank gilt allen Beteiligten, die an der Studie im Rahmen meiner Bachelorarbeit teilgenommen haben.

Danken möchte ich auch meinen Freunden, die mich während der gesamten Zeit moralisch unterstützt haben.

Inhaltsverzeichnis

Abbildungsverzeichnis	ix
Tabellenverzeichnis	xi
1 Einleitung	1
1.1 Motivation	1
1.2 Problemstellung	2
1.3 Aufbau und Struktur der Arbeit	2
2 Verwandte Arbeiten	3
3 Grundlagen	5
3.1 Online Diskussion	5
3.2 Argument und Argumentation	5
3.3 Relationen bei Argumenten	6
3.4 D-BAS: Dialog-basiertes Argumentationssystem	8
3.4.1 Einführung in D-BAS	8
3.4.2 Argumentationszyklus in D-BAS	9
3.5 Colour Coding	11
4 Pre-Test und Testdurchführung	13
4.1 Ziel der Studie	13
4.2 User Testing	13
4.3 Umfragebogen	14
4.4 Pre-Test	16
4.4.1 Auswertung	17
4.4.2 Änderungen durch den Pre-Test	18
4.5 Durchführung des ersten Tests	19
4.6 Durchführung des zweiten Tests	19
5 Evaluation	21

5.1	Auswertung des ersten Tests	21
5.2	Auswertung des zweiten Tests	22
5.3	Vergleich der Tests	23
5.4	Statistische Auswertung	25
5.5	Verbesserungen	26
6	Fazit und Ausblick	31
6.1	Fazit	31
6.2	Ausblick	32
	Literaturverzeichnis	53

Abbildungsverzeichnis

3.1	Support Schaubild	6
3.2	Rebut Schaubild	7
3.3	Undercut Schaubild	7
3.4	Undermine Schaubild	8
3.5	Argumentationszyklus	9
3.6	Dialogbeispiel D-BAS (1)	10
3.7	Dialogbeispiel D-BAS (2)	10
3.8	Dialogbeispiel D-BAS (3)	11
3.9	Dialogbeispiel D-BAS (4)	12
3.10	Dialogbeispiel D-BAS mit Colour Coding	12
4.1	Zusammenhang zwischen der Anzahl an Probanden und dem prozentualen Anteil ge- fundener Usability-Probleme [Rei06]	14
4.2	Auszug Pre-Test	18
4.3	Flyer zur Probandenwerbung	20
5.1	Beispiel mit möglichen Verbesserungen	29

Tabellenverzeichnis

4.1	Daten Pre-Test Rebut	17
4.2	Daten Pre-Test Undercut	17
4.3	Daten Pre-Test Undermine	18
5.1	Schema zur Auswertung des Feedbacks	21
5.2	Auswertungsdaten mit dem Verfahren aus Tabelle 5.1	22
5.3	Auswertungsdaten zweiter Test	23
5.4	Auswertungsdaten Vergleichsfragebogen	24
5.5	Direkter Vergleich Bogen Nr. 1/2	24
5.6	Vergleich Feedbackmöglichkeiten	25

Kapitel 1

Einleitung

1.1 Motivation

Um online Diskussionen zu führen, gibt es mittlerweile mehrere Möglichkeiten. Viele Internetnutzer verwenden Foren, Blogs oder auch Social-Media Seiten, um auf diese Weise zu diskutieren und verschiedenste Meinungen untereinander auszutauschen. Diese Möglichkeiten der online geführten Diskussionen sind oftmals unübersichtlich und unstrukturiert, insbesondere dann, wenn mehrere hundert Nutzer an diesen beteiligt sind. Hierbei ist es oftmals schwierig, innerhalb der Diskussion zu navigieren oder bestimmte Punkte zu finden. Die Interessen und Referenzen des Nutzers, die er schon beispielsweise in anderen Diskussionen geäußert hat, bleiben dabei unberücksichtigt. Die Aussagen mit Freitext sind hierbei oft unkoordiniert und ungeordnet, wobei oft allgemein bekannte Probleme wie Balkanisierung, Spam und Redundanz entstehen. Aus diesem Grund sind strukturierte Antwortmöglichkeiten innerhalb einer Diskussion sehr wichtig.

Genau mit dieser Problemstellung befasst sich D-BAS, ein neu entwickeltes, dialog-basiertes Argumentationssystem, welches die gerade genannten Probleme in Zukunft reformieren und die Navigation beziehungsweise Argumentation in online geführten Diskussionen für den Nutzer erheblich vereinfachen soll. Dies gilt es in dieser Arbeit zu evaluieren.

1.2 Problemstellung

D-BAS ermöglicht dem Benutzer¹, Diskussionen über das Internet zu führen. Da es dem Nutzer verschiedenste Möglichkeiten der Interaktion vorgibt, soll herausgefunden werden, ob D-BAS alle Möglichkeiten der Interaktion beinhaltet, die auch real in Diskussionen von Personen verwendet werden. Dazu stellt sich die Frage, wie genau ein Nutzer mit und ohne Vorgaben des Systems reagiert. Genau diese Unterschiede herauszufinden und somit dahingehende Verbesserungsvorschläge für D-BAS zu entwickeln ist das Ziel dieser Arbeit.

1.3 Aufbau und Struktur der Arbeit

Diese Arbeit ist in fünf Kapitel unterteilt. Das erste Kapitel befasst sich mit der Motivation und Problemstellung der Thematik sowie dem Aufbau und der Struktur der Arbeit. Das zweite Kapitel geht kurz auf verwandte Arbeiten ein, die bei der Erstellung der Bachelorthesis geholfen haben und an denen sich orientiert wurde. Im dritten Kapitel werden Grundlagen der Argumentationsthematik erläutert und an verschiedenen Beispielen erklärt, wichtige Begriffe definiert und es folgt eine Einführung in das dialog-basierte Argumentationssystem D-BAS. Das darauffolgende vierte Kapitel befasst sich mit der Testplanung und Durchführung, in welchem die Benutzerinteraktionen des Argumentationssystems D-BAS analysiert werden. Hierbei wird der Weg vom Pre-Test über den ersten Test bis zum Abschluss des zweiten Tests geschildert, vor dem einige Veränderungen am System durchgeführt wurden. Das fünfte Kapitel befasst sich danach mit der Evaluation. Hierbei werden die Studien aus dem vierten Kapitel analysiert und ausgewertet. Es folgt eine Bewertung der getesteten Benutzerinteraktionen des Argumentationssystems D-BAS. Außerdem werden Verbesserungsvorschläge für dieses System unterbreitet. Schlussendlich wird im sechsten Kapitel eine Zusammenfassung der Arbeit, ein Ausblick zur weiteren Optimierung des Systems beschrieben und es folgt das Literaturverzeichnis. Im Anhang befinden sich die Bögen aller Tests als auch eine CD mit allen Unterlagen, welche die Testerbenisse sowie eine digitale Version der Umfragebögen und einer elektronischen Version dieser Arbeit beinhaltet.

¹ Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten gleichwohl für beide Geschlechter.

Kapitel 2

Verwandte Arbeiten

Diese Bachelorthesis befasst sich mit der Analyse der Feedbackmöglichkeiten von Benutzern mit einem dialog-basierten Argumentationssystem. Die Thesis orientiert sie sich an anderen Arbeiten, die sich mit Usability (Benutzerfreundlichkeit) beschäftigen.

Vor der Einarbeitung in das eigentliche Thema dieser Arbeit wurde eine Bachelorthesis zum Thema „Usability-Untersuchung von Liquid Democracy-Systemen“ durchgearbeitet. In dieser Arbeit wird die grundlegende Thematik zum Thema Usability-Test beschrieben. An wichtigen Punkten zur Planung, Ausarbeitung und Durchführung eines solchen Testes wurde sich in dieser Arbeit orientiert [Erd13].

Des Weiteren wurde eine Diplomarbeit zum Thema „Konzeption, Entwicklung und Usability Evaluation einer Webanwendung für die Verwaltung von Webhosting Leistungen“ gelesen, welche sich mit der Gebrauchstauglichkeit von Webseiten und Webanwendungen beschäftigt. Ebenfalls werden die Methoden der Usability Evaluation behandelt, an denen sich diese Arbeit im weiteren Verlauf orientiert [Rei06].

Kapitel 3

Grundlagen

3.1 Online Diskussion

Im Internet besteht die Möglichkeit, durch Online-Partizipationen gegenseitig zu diskutieren, unabhängig von Zeit und Ort sowie der Anzahl an Diskussionsteilnehmern. Schaut man auf die verschiedenen Möglichkeiten, wie beispielsweise Blogs, Social-Media oder auch Internetforen, so wird deutlich, dass der Austausch von Meinungen hierbei oft alles andere als rational ist. Foren beispielsweise sortieren nicht nach der Anzahl der Teilnehmer, sondern nach Beiträgen. Oftmals fehlen Struktur und Übersichtlichkeit, da vieles wiederholt wird oder es zu ähnlichen Meinungen von Benutzern kommt [KBBM16, S.1-2].

3.2 Argument und Argumentation

Jede Diskussion befasst sich mit einem **Thema**. Dieses kann zum Beispiel „*Elektroautos - Autos der Zukunft?*“ sein. **Statements** sind einfache Aussagen, die in einer Diskussion verwendet werden, wie beispielsweise „*E-Autos sind optimal für den Stadtverkehr*“ oder „*E-Autos produzieren keine stinkenden Abgase*“. Teilnehmer können diese als wahr oder falsch betrachten. Des Weiteren gibt es richtungsweisende Aussagen, welche als **Position** bezeichnet werden und eine Handlung nach sich ziehen. „*In der Stadt sollten nur noch elektronisch betriebene KFZs zugelassen sein*“ ist ein Beispiel für eine Position, da hiermit verbunden ist, dass mit Kraftstoff betriebenen KFZs in der Stadt verboten werden sollen [KBBM16, S. 6].

In einer Diskussion werden Argumente miteinander verknüpft. Ein **Argument** besteht dabei immer aus mehreren Statements, wobei man zwischen **Konklusion** und **Prämisse** unterscheidet. Die Kon-

klusion ist das Statement, welches begründet werden soll. Die Prämisse soll dabei die Konklusion stützen oder kann diese angreifen. Für jede Konklusion kann es eine oder mehrere Prämissen geben, denn es können natürlich mehrere Statements unterstützend auf die Konklusion wirken [Bay07, S. 18-19]. Am besten ist dies Anhand eines Beispiels zu erklären:

1. Prämisse: *E-Autos sind optimal für den Stadtverkehr.*

2. Prämisse: *Modell XY ist ein E-Auto.*

Konklusion: *Modell XY ist optimal für den Stadtverkehr.*

Beide Prämissen unterstützen hierbei dieselbe Konklusion. Es wird argumentiert, um sich zu rechtfertigen oder Behauptungen zu begründen. Mit diesen Argumenten können Diskussionen geführt werden, um entweder gegeneinander oder auch miteinander verschiedene Argumente für oder gegen eine Behauptung zu finden [Bay07, S. 18-19].

3.3 Relationen bei Argumenten

Im nachfolgenden Abschnitt werden die Reaktionsmöglichkeiten bezüglich einer Argumentation beschrieben. Im folgenden wird (+) verwendet, wenn ein Statement oder Argument unterstützend und (-), wenn es angreifend wirkt.

Support Ein **Support** bezeichnet zwei Argumente mit gleicher Konklusion. Hierzu wird das Beispiel in Abbildung 3.1 betrachtet. Sowohl das erste Argument mit der Prämisse „E-Autos verursachen

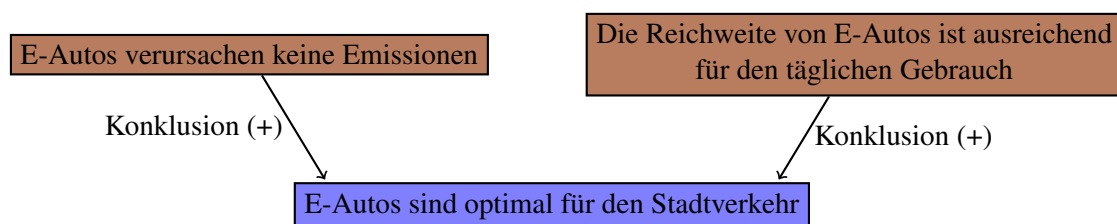


Abbildung 3.1: Beispiel eines Supports innerhalb des Arguments

keine Emissionen“ als auch das zweite Argument mit der Prämisse „Die Reichweite von E-Autos ist ausreichend für den täglichen Gebrauch“ besitzen die selbe Konklusion „E-Autos sind optimal für den Stadtverkehr“. Beide Argumente unterstützen sich gegenseitig.

Rebut Der **Rebut** bezeichnet ein Argument, welches die negierte Konklusion eines anderen Arguments ist, welche genau diese angreift. [KBBM16, S.5] Im Beispiel aus Abbildung 3.2 besteht das Argument aus der Konklusion „E-Autos sind optimal für den Stadtverkehr“ und der Prämisse „E-Autos produzieren keine stinkenden Abgase“. Die Prämisse des zweiten Arguments „Die Reichweite ist

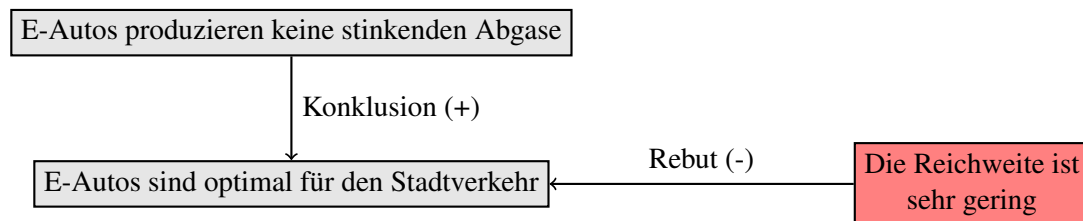


Abbildung 3.2: Angriff der Konklusion durch den Rebut

sehr gering“ wirkt somit schwächend auf die Konklusion des ursprünglichen Arguments [KBBM16, S.5].

Undercut Unter einem **Undercut** versteht man die angreifende Wirkung eines Arguments auf die Relation zwischen Prämisse und Konklusion eines anderen Arguments. Betrachtet wird nun das Argument mit Prämisse „Die Reichweite von E-Autos ist ausreichend für einen Tag“ sowie der Konklusion „E-Autos sind optimal für den Stadtverkehr“ aus Abbildung 3.3. Hier wirkt das Argument „Fahrräder

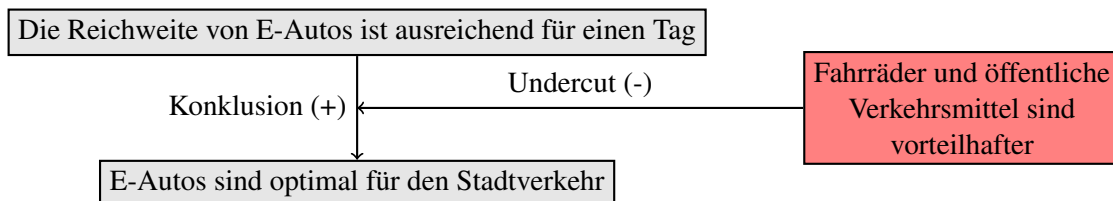


Abbildung 3.3: Angriff auf die Relation mittels Undercut

und öffentliche Verkehrsmittel sind vorteilhafter“ angreifend auf die Relation zwischen Prämisse und Konklusion des ursprünglichen Arguments [KBBM16, S. 5].

Undermine Der **Undermine** bewirkt die Schwächung eines Arguments durch die Prämisse eines anderen Arguments. Im Beispiel aus Abbildung 3.4 wird die Prämisse „Die Reichweite von E-Autos beträgt 300 Km“ durch das Argument „Die Leistung der Batterie pro Ladeinheit und somit auch die maximale Distanz der E-Autos ist unterschiedlich“ geschwächt [KBBM16, S. 5].

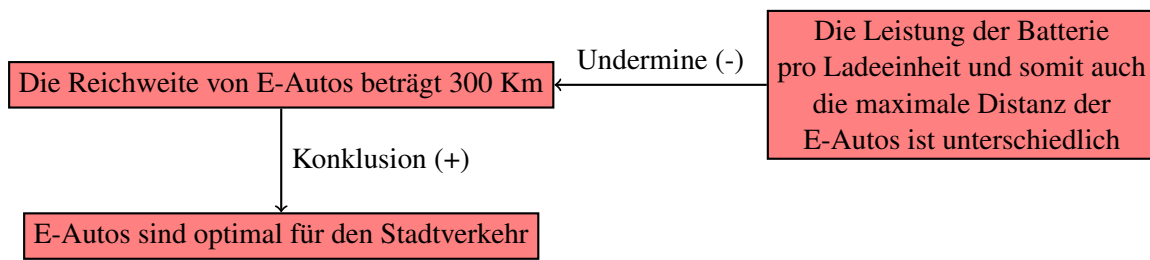


Abbildung 3.4: Schwächung der Prämisse durch den Undermine

3.4 D-BAS: Dialog-basiertes Argumentationssystem

Dialog-basierte Online-Argumentationen sollen jedem Benutzer ermöglichen, an einer online geführten Argumentation teilzunehmen, ohne dass dieser Vorwissen in diesem Bereich hat. Der Benutzer soll in der Lage sein, verschiedene Möglichkeiten aus einer Auswahl an Alternativen zu wählen oder auch selber eine Aussage zu schreiben. Basierend auf den Antworten der anderen Benutzer und den gesammelten Daten wählt das System das nächste Argument aus, welches dem Benutzer angezeigt wird. Durch diese Vorgehensweise wird der Benutzer vom System durch die Argumentation geleitet und es werden ihm die für ihn wahrscheinlich interessantesten Argumente vom System angezeigt. Außerdem bietet es Feedbackmöglichkeiten zu diesen Argumenten. Ein Benutzer, der hierbei eine bestimmte Position einnimmt, kann somit vom System speziell mit bestimmten Argumenten konfrontiert werden. Soziale Einflüsse sowie der *follow the herd*-Effekt¹ können zunächst vermieden werden und die Argumente des Systems können dem Benutzer somit neutral präsentiert werden. Das System aktualisiert hierbei mit jeder Auswahl oder Antwort des Benutzers seine Wissensbasis.

3.4.1 Einführung in D-BAS

D-BAS bietet dem Benutzer eine Benutzerverwaltung, welche die Registrierung, ein Login sowie das Verwalten von persönlichen Einstellungen beinhaltet. Außerdem besitzt das System die Möglichkeit, URLs (Uniform Resource Locator) der Diskussionen mit anderen Internetnutzern per sozialen Medien wie Facebook, Twitter oder Google+ zu teilen. Registrierte Nutzer können außerdem jederzeit Aussagen hinzufügen oder vorhandene Aussagen bearbeiten. Dies kann beispielsweise dann der Fall sein, wenn jemand eine Aussage für ungeeignet empfindet. Anonyme Benutzer, welche nicht registriert sind, können die Diskussion ebenfalls führen, jedoch ist es ihnen nicht möglich, Aussagen zu bearbeiten oder hinzuzufügen. D-BAS bietet außerdem die Möglichkeit, in einer Diskussion schrittweise zurückzugehen, um mögliche andere Argumente auszuwählen. Von registrierten Nutzern wird

¹etwas tun, was die meisten Anderen auch tun

der Verlauf in Form einer History gespeichert, in dem diese die Möglichkeit haben, auch in älteren Diskussionen an verschiedenen Punkten erneut in diese einzusteigen. Um die Oberfläche von D-BAS möglichst benutzerfreundlich und übersichtlich zu halten, wurde diese in Form von Sprechblasen gestaltet, ähnlich der modernen Chat-Clients wie WhatsApp oder Facebook Messenger. Der Nutzer kann somit durch einfaches Klicken auf eine Blase zum gewählten Punkt in der Diskussion navigieren [KBBM16].

3.4.2 Argumentationszyklus in D-BAS

Ein wichtiger Punkt ist die Auswahl der nächsten Argumente, welche dem Benutzer angezeigt werden. Hierbei besteht jede Navigation innerhalb einer Diskussion aus drei Schritten:

1. Auswahl
2. Konfrontation
3. Feedback

Zur Verdeutlichung dient die nachfolgende Abbildung. Der Benutzer wählt eine Aussage, die für

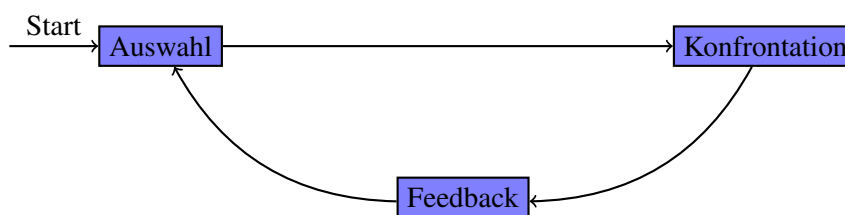


Abbildung 3.5: Argumentationszyklus in D-BAS

ihn Interessant ist. Daraufhin wird er vom System gefragt, ob er dieser Aussage zustimmt oder diese ablehnt. Alternativ kann er sich eine weitere Aussage für seine eben gewählte anzeigen lassen, wenn er sich nicht sicher ist. Im nächsten Schritt kann der Nutzer seinen wichtigsten Grund für diese Aussage auswählen. Nun wird dieser durch das System mit einem Gegenargument konfrontiert (An dieser Stelle greift D-BAS mittels Undercut, Undermine oder Rebut an, siehe Kapitel 3.3). Hat er dies getan, folgt im nächsten Schritt die Auswahlmöglichkeiten des Feedbacks vom Nutzer an das System, welche wie folgt gegliedert sind:

- Die angreifende Prämisse ist falsch
- Die angreifende Prämisse ist richtig, damit auch die daraus folgende Konklusion
- Die angreifende Prämisse und negierte Konklusion ist richtig, jedoch begründet die Prämisse die negierte Konklusion nicht
- Das angreifende Argument wird akzeptiert, jedoch hat der Benutzer eine für sich bessere, weitere Prämisse für seine Konklusion
- Ein anderes, angreifendes Argument anzeigen

Um die Argumentationsmöglichkeiten zu verdeutlichen, wird nun ein Dialogbeispiel betrachtet, welches mit D-BAS durchgeführt wurde. In Abbildung 3.6 wird der Nutzer nach der eigenen Interesse

Ich möchte über die Aussage reden, dass ...

- ☐ E-Autos sind sehr teuer, rentieren sich niemals in der Anschaffung und sind daher nur was für Idealisten mit zu viel Geld.
- ☐ E-Autos keine Emissionen verursachen
- ☐ Elektroautos sehr günstig im Unterhalt sind
- ☐ E-Autos optimal für den Stadtverkehr sind
- ☐ Um eine neue Aussage hinzuzufügen, klicken Sie bitte hier um sich anzumelden.

Abbildung 3.6: Dialogbeispiel D-BAS (1)

gefragt. Durch Auswahl oder dem Einfügen einer Möglichkeit beginnt dieser mit der Diskussion (siehe Kapitel 3.4.2). Im Verlauf dieses Dialogbeispiels interessiert sich der Nutzer dafür, dass „E-Autos optimal für den Stadtverkehr sind“. Nun befindet er sich im zweiten Schritt (Abbildung 3.7), indem

Was halten Sie davon, dass E-Autos optimal für den Stadtverkehr sind?

- ☐ Ich **akzeptiere** die Aussage, dass E-Autos optimal für den Stadtverkehr sind
- ☐ Ich **widerspreche** der Aussage, dass E-Autos optimal für den Stadtverkehr sind
- ☐ Ich **weiß es nicht**. Zeige mir eine Aussage dafür, dass E-Autos optimal für den Stadtverkehr sind

Abbildung 3.7: Dialogbeispiel D-BAS (2)

er diesen Standpunkt unterstützen, ablehnen oder sich eine Aussage zur Unterstützung anzeigen lassen kann (siehe Kapitel 3.4.2). In diesem Beispiel akzeptiert der Nutzer die Aussage, welche dadurch

nun seinen Standpunkt widerspiegelt. Nach der Auswahl wird der Nutzer nach der Begründung für

Ich bin der Meinung, dass E-Autos optimal für den Stadtverkehr sind.
11 weitere Teilnehmer/innen mit dieser Meinung.

Jetzt

Was ist Ihr wichtigster Grund dafür, dass E-Autos optimal für den Stadtverkehr sind, richtig ist?
Weil...

☐ die Reichweite von Elektroautos ausreichend für mindestens 300km ist

☐ Sofern Sie eine neue Aussage hinzufügen möchten, klicken Sie bitte hier um sich anzumelden.

Abbildung 3.8: Dialogbeispiel D-BAS (3)

seine Auswahl gefragt (Abbildung 3.8). Hierbei kann er zwischen der vorgegebenen Aussage „*Die Reichweite von E-Autos ist ausreichend für mindestens 300 Km*“ und der Möglichkeit, eine eigene Aussage hinzuzufügen, wählen. Im Beispiel entscheidet er sich für die vorgegebene Aussage. Nun präsentiert D-BAS ein Argument gegen die Konklusion „*Die Reichweite von E-Autos ist ausreichend für mindestens 300 Km*“ des Argumentes „*E-Autos sind optimal für den Stadtverkehr*“. Über dieses Gegenargument „*Die Ladezeit der Batterien kann bis zu 12 h dauern und man hat insbesondere tagsüber diese Zeit oftmals nicht*“ kann der Nutzer nun hinsichtlich seiner Meinung entscheiden. Hierbei bietet D-BAS ihm die in Abbildung 3.9 gezeigten Möglichkeiten. Je nach Auswahl des Nutzers gibt D-BAS ihm ein neues Argument vor, bei dem dieser die gleichen Auswahlmöglichkeiten, jedoch nun auf das neue Argument bezogen, präsentiert bekommt. Dies wird so lange fortgesetzt, bis kein neues Argument mehr in der Datenbank verfügbar ist.

3.5 Colour Coding

Während der Durchführung wurden bereits Veränderungen von den Entwicklern am System durchgeführt. Um die Auswahlmöglichkeiten des Systems für den Nutzer verständlicher zu gestalten, wurden die verschiedenen Standpunkte und Argumente farblich markiert und durch Angaben wie „*deren Argument*“, „*deren Standpunkt*“ und „*mein Standpunkt*“ ersetzt (siehe Abbildung 3.10). Im Vergleich mit Abbildung 3.9 sind die Auswahlmöglichkeiten des Nutzers nun kürzer und es wird durch die farblichen Markierungen sofort deutlich, welches Argument oder welcher Standpunkt sich worauf bezieht. Dies erleichtert die Interaktion des Benutzers mit dem System.

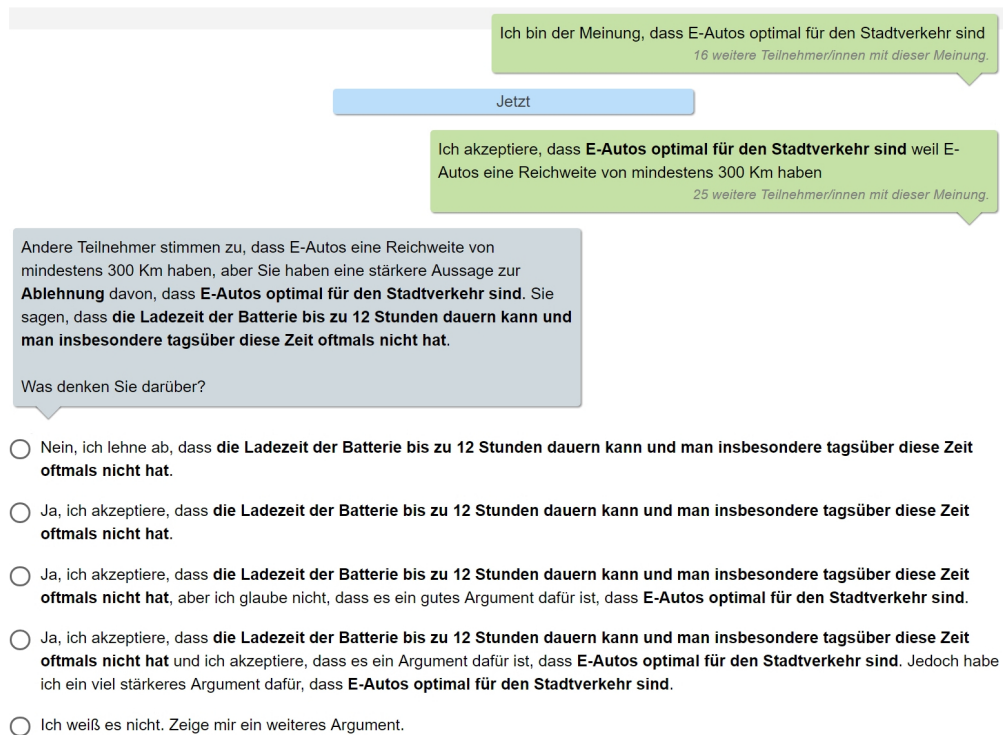


Abbildung 3.9: Dialogbeispiel D-BAS (4)



Abbildung 3.10: Dialogbeispiel D-BAS mit Colour Coding

Kapitel 4

Pre-Test und Testdurchführung

Im folgenden Kapitel wird die Planung und Durchführung des Pre-Tests sowie der beiden Haupttests beschrieben. Hierbei werden alle Planungsabschnitte von der Überlegung der Zielsetzung über die Durchführung und Auswertung des Pre-Tests bis hin zur Erstellung der Umfragebögen geschildert. Anschließend erfolgt die Beschreibung der Testdurchführung.

4.1 Ziel der Studie

In der Studie im Rahmen dieser Bachelorthesis sollte herausgefunden werden, wie die Benutzer auf Situationen, welche mit D-BAS erstellt und vorgegeben wurden, reagieren. Hierzu wurde das Thema *Elektroautos* gewählt. Die Datenbank wurde ergänzt, sodass mit verschiedenen Möglichkeiten auf die Interaktionen des Benutzers reagiert werden kann. Dieser sollte zuerst seine eigene Meinung äußern, bevor D-BAS ihm Antwortmöglichkeiten zur Interaktion vorgab. Ziel war, herauszufinden, ob der Benutzer mit vorgegebenen Auswahlmöglichkeiten des Systems genauso handelt wie wenn er alleine entscheiden und seine Meinung zur vorgegebenen Diskussion äußern soll. Im Fokus standen die Feedbackmöglichkeiten des Nutzers mit dem System.

4.2 User Testing

Der Nutzertest ist eine komplexe Methode, die angewendet wird, um einen Prototypen, der von Nutzern bedient werden kann, zu testen. Hierbei hat dieser folgende Kennzeichen:

1. Usability des Prototyps und dessen Entwicklungsprozess soll verbessert werden
 2. Die Probanden sind Personen aus einer gewählten Zielgruppe
 3. Die Probanden führen realistische Aufgaben aus
 4. Das individuelle Verhalten der Teilnehmer beim Umgang mit dem Prototypen wird beobachtet
 5. Die Ergebnisse werden analysiert und Änderungen zur Beseitigung der Probleme vorgeschlagen
- [Rei06]

Auswahl der Probanden Bei der Auswahl der Probanden ist zu beachten, dass diese aus einer gewählten Zielgruppe stammen. Hierbei ist wichtig, dass die gewählten Probanden das System später auch nutzen werden. Oftmals kommen daher Entwickler oder Personen, die in irgend einer Art und Weise am Prototypen mitgearbeitet haben, nicht in Frage [Rei06, S.39].

Anzahl der Probanden Die minimale Gruppengröße an Probanden bei einem Usability-Test beträgt vier bis fünf Personen, da diese bereits ungefähr 75% der Probleme im System finden. Eine Vergrößerung der Probandengruppe würde wenige Prozentpunkte gutmachen (siehe Abbildung 4.1). Die Probandengruppe kann in kleinere Untergruppen eingeteilt werden, welche nach Alter oder Ge-

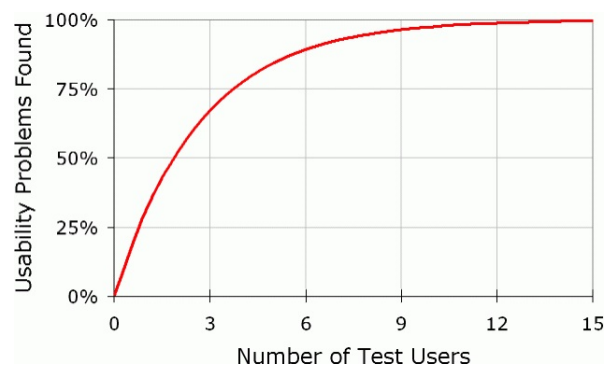


Abbildung 4.1: Zusammenhang zwischen der Anzahl an Probanden und dem prozentualen Anteil gefundener Usability-Probleme [Rei06]

schlecht gebildet werden können [Rei06]. Für die Studie, welche in dieser Arbeit behandelt wird, wurden drei Gruppen mit jeweils sechs Probanden ausgewählt.

4.3 Umfragebogen

Bei der Erstellung der Umfragebögen musste folgendes beachtet werden:

D-BAS benutzt drei Möglichkeiten, den Standpunkt des Nutzers anzugreifen (Undercut, Rebut und Undermine, siehe Kapitel 3.3). Um die Testpersonen nicht zu überfordern und um zu erreichen, dass sich diese unvoreingenommen auf ihre persönliche Meinungen konzentrieren, wurde für jede der drei Interaktionsmöglichkeiten des Systems ein einzelner Umfragebogen erstellt. Dies stellte sicher, dass keine der Testpersonen vorher einen ähnlichen Umfragebogen bearbeitet hat oder schon vorab wusste, worüber diese Studie handelt. Der Umfragebogen sollte kurz und leicht verständlich sein, da D-BAS mit Personen getestet werden sollte, die das System zum ersten Mal benutzen und kein besonderes Vorwissen im Bereich der Argumentation haben. Dies ist wichtig, da das System später für Benutzer aus jeder Zielgruppe verständlich sein soll. Die nächste Überlegung war, welche statistischen Daten der Testpersonen für diese Studie überhaupt interessant sind. Da, wie schon erwähnt, die Studie keine bestimmte Zielgruppe voraussetzt, wurden im Umfragebogen lediglich Geschlecht, Alter und Bildungsgrad abgefragt. Nach Abschluss dieser Überlegungen wurden die ersten Fassungen der Umfragebögen erstellt. Hierzu wurde der Anbieter Umfrageonline (www.umfrageonline.com) gewählt, welcher Studenten, die an einer Hochschule eingeschrieben sind, seinen Umfragebaukasten kostenfrei zur Verfügung stellt. Mit Hilfe dieses Anbieters kann online ein Umfragebogen erstellt werden, indem man jede Seite nach dem Baukastenprinzip gestalten kann und vorhandene Tools, welche dort zur Verfügung stehen, in seine Umfrage einbaut. Hierbei bietet Umfrageonline verschiedene Möglichkeiten, wie beispielsweise eine Frage mit Antwortmöglichkeiten, Bewertungsskalen oder auch leere Textfelder, indem die Testperson seine Antwort als ausgeschriebenen Text eintippen soll. Ein Vorteil dieses Tools ist, dass es ohne Werbung arbeitet und die Testergebnisse als Microsoft Excel Datei (XLS) oder Comma-separated values Datei (CSV) exportiert werden können. Dies erleichtert die Analyse der Auswertungsergebnisse, da keine manuelle Eingabe dieser erforderlich ist, um sie zu verwerten.

Schema Umfragebogen Nach reichlicher Überlegung ergab sich für die Gliederung des Umfragebogens folgendes Schema:

1. Begrüßung und Erläuterung des Umfragebogens
2. Datenschutzerklärung
3. Erfassung der gewählten persönlichen Daten (Geschlecht, Alter, Bildungsgrad)
4. Abbildung 1 mit einem Screenshot von D-BAS und leerem Textfeld mit der Aufforderung an die Testperson, ihre eigene Meinung zur vorgegebenen Diskussion zu erläutern
5. Abbildung 2 (welche aus dem gleichen Screenshot bestand) mit Antwortmöglichkeiten, die von D-BAS zur Verfügung gestellt werden. Hierbei wurde die Möglichkeit ergänzt, dass keine der angegebenen Antworten zutrifft.

4.4 Pre-Test

Nach der Erstellung der Umfragebögen war es nun wichtig herauszufinden, ob diese auch ihren Zweck erfüllen und für die Testpersonen überhaupt verständlich sind. Dazu bot es sich an, einen Pre-Test durchzuführen. Ein Pre-Test ist ein wichtiger Teil einer Studie, welcher mit einem Umfragebogen durchgeführt wird. Durch ihn sollen Verständnisprobleme, Ungenauigkeiten und Fehler im eigentlichen Test aufgedeckt werden. Dies ist wichtig, da Fehler, die beispielsweise nach der Durchführung der Studie entdeckt werden, dazu führen können, dass diese (oder auch ein Teil der Studie) unbrauchbar ist. Hierbei ist zu beachten, dass die Fragen so formuliert sind, dass die Probanden diese auch richtig verstehen. Gerade missverständlichen Formulierungen können oftmals erst beim Pre-Test erkannt und behoben werden. Ein weiterer Vorteil liegt darin, dass eventuell unnötige Fragen, die gar nicht zum Ziel der Studie beitragen, gestrichen werden können. Hilfreich ist es, mit den Probanden des Pre-Tests über Fragestellung und Verständlichkeit zu diskutieren, um eindeutig sicherzustellen, dass diese so verstanden werden, wie es intendiert ist. Der Pre-Test wird mit einer deutlich kleineren Gruppe an Probanden durchgeführt. Allerdings sollte beachtet werden, dass die Bedingungen des richtigen Tests so gut wie möglich simuliert werden. Vorteilhafterweise wird ein Teil der Probanden so gewählt, dass dieser aus Fachleuten oder Personen besteht, die sich mit dem Thema schon einmal auseinander gesetzt haben [Kol12, S. 38-39].

Natürlich sollte der Anteil an Fachleuten nicht zu groß sein, da der richtige Test dann mit Probanden jeglicher Zielgruppe durchgeführt wird. Für den Pre-Test dieser Studie wurden die mit Umfrageonline erstellten Umfragebögen genommen. Diese wurden jedoch so verändert, dass am Ende noch zwei Seiten zur Bewertung angehängt wurden. Hierbei sollte auf einer Likert-Skala abgestimmt werden, welche die Zustimmung oder Ablehnung der Probanden auf einer Skala von **1: trifft gar nicht zu** bis **5: trifft voll und ganz zu** misst [Kol12, S. 27]. Folgende Punkte sollten bewertet werden:

- Bewertungsskala von 1-5 zu den Punkten Verständlichkeit und Übersichtlichkeit
- Kritik und Anregungen der Probanden (hierzu wurde ein leeres Textfeld zur Verfügung gestellt)

Des Weiteren sollte der Pre-Test (wie auch später für den Haupttest angedacht) als Thinking-Aloud-Test¹ durchgeführt werden, damit Probleme bezüglich des Verständnisses sowie Anregungen und Kritik direkt notiert werden konnten.

Der Pre-Test wurde mit sechs Personen durchgeführt, wobei drei an D-BAS mitgearbeitet haben oder sich in dem Themengebiet auskennen (Fachleute) und drei Studenten, denen weder D-BAS bekannt

¹Testpersonen sollen während der Beantwortung der Fragen ihre Gedanken, Reaktionen und Handlungen laut aussprechen. [Erd13] Der anwesende Testleiter kann somit dem Probanden bei eventuell auftretenden Fragen oder Unklarheiten weiterhelfen.

war noch sich mit dem Thema „dialog-basierte Argumentation“ auseinandergesetzt haben. Für jeden der drei Umfragebögen (Undercut, Rebut, Undermine) wurde nun zufällig jeweils eine Person der Fachleute sowie ein Student gewählt, sodass pro Umfragebogen des Pre-Tests jeweils zwei Probanden zur Verfügung standen. Der Pre-Test wurde an einem Vormittag mit den sechs Probanden im Gebäude 25.12 der Heinrich-Heine-Universität in Düsseldorf durchgeführt.

4.4.1 Auswertung

Um den Pre-Test auszuwerten, werden zuerst die Daten betrachtet, welche bei der Bewertung bezüglich Verständlichkeit und Übersichtlichkeit gesammelt wurden. Um einen gemittelten Wert zu erhalten, wird für die Datenauswertung das arithmetische Mittel $\bar{x}$ berechnet, welches wie folgt definiert ist:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i, n \in \mathbb{N}^{>0}$$

mit n Teilnehmern und der jeweiligen Bewertung x_1 bis x_n
[BMSMS03, S. 38]

Dazu werden die drei verschiedenen Testergebnisse einzeln betrachtet (im Folgenden ist $n = 2$).

Daten Pre-Test Rebut	Proband 1	Proband 2	$\bar{x}$
Die Grafiken waren verständlich	4	5	4,5
Die Frageseiten waren übersichtlich	4	5	4,5
Ich hatte keine Verständnisprobleme	4	4	4

Tabelle 4.1: Daten Pre-Test Rebut

Daten Pre-Test Undercut	Proband 1	Proband 2	$\bar{x}$
Die Grafiken waren verständlich	5	4	4,5
Die Frageseiten waren übersichtlich	5	3	4
Ich hatte keine Verständnisprobleme	3	1	2

Tabelle 4.2: Daten Pre-Test Undercut

Bei Betrachtung der arithmetischen Mittel fällt der Wert zur Aussage „Ich hatte Verständnisprobleme“ auf (besonderes in Tabelle 4.2). Dies wurde auch bereits durch Äußerungen der Probanden während des Tests geäußert. Viele waren unsicher und wussten nicht genau, was von ihnen verlangt wurde. Oft wurde geäußert, dass nicht klar sei, ob die eigene Meinung über den Aufbau des Systems, welchen

man anhand des Screenshots erkennen konnte, oder der Bezug auf den inhaltlichen Kontext verlangt war.

Daten Pre-Test Undermine	Proband 1	Proband 2	$\bar{x}$
Die Grafiken waren verständlich	3	4	3,5
Die Frageseiten waren übersichtlich	5	4	4,5
Ich hatte keine Verständnisprobleme	4	4	4

Tabelle 4.3: Daten Pre-Test Undermine

4.4.2 Änderungen durch den Pre-Test

Um den vierten und fünften Punkt (siehe Aufzählung im Abschnitt „Schema Umfragebogen“ in Kapitel 4.3) verständlicher zu gestalten und die Probleme, die durch den Pre-Test aufgedeckt wurden, zu verdeutlichen, wird auf diese im folgenden Abschnitt kurz eingegangen. Betrachtet wird nun die selbe Abbildung, die auch einigen Probanden gezeigt wurde (Abbildung 4.2). Hier zeigten sich die Proban-

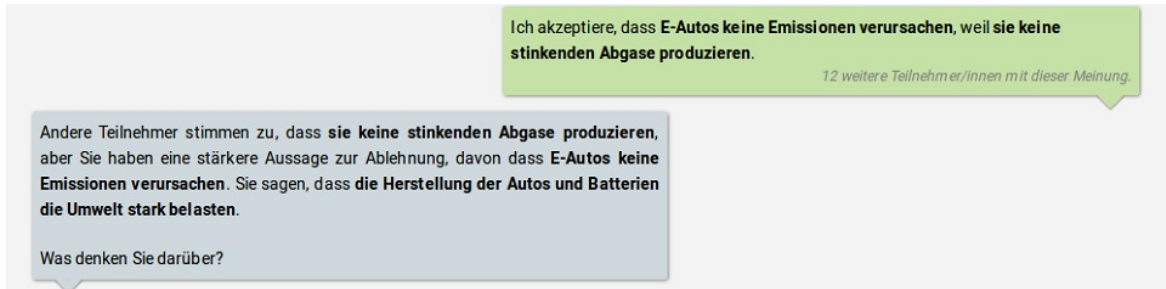


Abbildung 4.2: Auszug Pre-Test

den oft ratlos und wussten nicht, was von ihnen verlangt war. Manche wollten die Aufmachung und Anordnung der Blasen beschreiben, andere den Inhalt. Da hier der inhaltliche Aspekt verlangt war, wurde die Fragestellung „Was denkst du darüber?“ geändert zu „Was ist deine Meinung zur Aussage der anderen Teilnehmer?“. Außerdem wurde vorab eine Erklärung eingefügt, dass es sich bei der nachfolgenden Abbildung (siehe Abbildung 4.2) um eine angefangene Diskussion über E-Autos handelt, welche mit D-BAS geführt wurde. Der Proband solle sich hierzu die bisherige Konversation durchlesen und seine Meinung dazu schildern. Durch diese Änderung war sichergestellt, dass dieser nun versteht, dass es sich um den Inhalt der Diskussion handele und nicht um optische Punkte, die bewertet werden sollten.

4.5 Durchführung des ersten Tests

Für die Durchführung der Umfrage wurden Personen aus den verschiedensten Zielgruppen (bezüglich Alter und Bildungsgrad) gesucht, um ein möglichst repräsentatives Ergebnis im Bezug auf den normalen Internetnutzer zu bekommen, wobei wie schon vorher erwähnt, es sonst keine besonderen Kriterien gab. Es war jedoch zu beachten, dass mögliche Probanden das Internet über Computer und Smartphone nutzen. Wichtig war außerdem, dass unter den Testpersonen ungefähr gleich viele männliche und weibliche, sowie Probanden aus jeder Altersgruppe, vertreten waren. Um den Testpersonen die Durchführung der Studie zu erleichtern, wurde diese bei der jeweiligen Person zu Hause durchgeführt. Dazu wurde mit jedem ein individueller Termin vereinbart, an dem der Proband zu Hause besucht und der Test am Laptop durchgeführt wurde. Hierbei wurde sichergestellt, dass die Bedingungen für alle Probanden gleich waren. Der jeweilige Test nahm pro Person ungefähr fünf bis zehn Minuten in Anspruch, je nach dem, wie schnell verstanden wurde, was gefragt war und wie ausführlich der Proband seine Antwort fasste. Da die Studie drei verschiedene Testbögen beinhaltete, wurden insgesamt achtzehn Probanden benötigt, jeweils sechs Probanden pro Bogen (Auswahl der Anzahl an Probanden siehe Kapitel 4.2).

4.6 Durchführung des zweiten Tests

Durch die Einführung des Colour-Coding (siehe Kapitel 3.5) stellte sich die Frage, ob diese Änderung einen Einfluss auf das Feedback der einzelnen Probanden hat, da sie die Länge der einzelnen Sätze verkürzt und zu einem besseren Verständnis führen soll. Zusätzlich wurden dazu kleine Veränderungen an den Formulierungen der einzelnen Aussagen des Systems vorgenommen. Um diese Änderungen zu testen und um überhaupt herauszufinden, ob diese zum besseren Verständnis des Nutzers beitragen, wurde ein zweiter Test geplant. Neben dem Colour-Coding und der Umformulierung mancher Statements wurde außerdem die erneute Befragung von sechs Testpersonen beschlossen, welche bereits am ersten Test teilgenommen haben. Diese zwei Gruppen an Probanden (erneut achtzehn Probanden, die noch nicht an dem Test teilgenommen haben sowie sechs Probanden, die auch schon am ersten Test teilgenommen haben) sollten den gleichen Test bekommen, in dem jedoch die oben genannten Änderungen eingebaut wurden. Jedoch bekamen die sechs Probanden, die schon teilgenommen haben, im Anschluss einen kurzen Bewertungsbogen, in dem Übersichtlichkeit und Verständlichkeit der beiden Fragebögen verglichen werden sollten. Hierbei sollte herausgefunden werden, ob die Veränderungen dem Probanden positiv auffielen. Die sechs Probanden wurden erneut nach Terminabsprache zuhause besucht, um den Test sowie den Fragebogen mit ihnen durchzuführen. Die neue Probandengruppe wurde genau wie die des ersten Tests ohne Vorwissen zur Bearbeitung an den Test gelassen. Um in geraumer Zeit neue achtzehn Probanden zu finden, wurde auf dem Campus der Heinrich-Heine-

Universität nach Freiwilligen gesucht. Vom Institut wurde hierfür eine Vergütung von 5 Euro pro Proband zur Verfügung gestellt, um die Attraktivität der Teilnahme zu steigern (siehe Abbildung 4.3), was die Suche erheblich vereinfachte. Erneut wurde jeweils ein individueller Termin vereinbart. Der



Abbildung 4.3: Flyer zur Probandenwerbung

Test wurde an den jeweiligen Terminen in Raum 25.12.02.47 der Universität durchgeführt. Der Ablauf des zweiten Tests war identisch mit dem des Ersten. Der jeweilige Proband sollte den Fragebogen am Computer beantworten, während erneut der Testleiter während der Beantwortung anwesend war. Bei der Lösung der eventuell entstehenden Probleme zum Verständnis konnte dieser dem Probanden während der Durchführung weiterhelfen. Für jeden der Probanden fanden sich die selben Testbedingungen vor.

Kapitel 5

Evaluation

5.1 Auswertung des ersten Tests

Um den Test auszuwerten, wurden die Eingaben der einzelnen Probanden, die diese eigenständig zu der angefangenen Diskussion eingegeben haben mit der jeweiligen Auswahl der Testperson der vorgegebenen Möglichkeiten von D-BAS verglichen.

Untersuchung des Feedbacks der Probanden Um eine Übersicht zu bekommen, wurde folgendes Verfahren aus Tabelle 5.1 gewählt. Hierbei wurden die Aussagen der einzelnen Probanden so gut

Auswertungsmethode	
0	Die eigene Aussage des Probanden stimmt mit der gewählten Auswahlmöglichkeit fast oder genau überein.
1	Die eigene Aussage des Probanden stimmt mit der gewählten Auswahlmöglichkeit nicht überein.

Tabelle 5.1: Schema zur Auswertung des Feedbacks

wie möglich mit den vorgegebenen Möglichkeiten verglichen, wobei beachtet wurde, dass die eigene Aussage nicht genau mit der gewählten Möglichkeit übereinstimmen kann. Es genügte, wenn sich die hauptsächliche Aussage des Probanden mit der Aussage der gewählten Möglichkeit des Systems deckte. War dies der Fall, wurde eine **0** notiert. Deckte sich die Aussage des Probanden überhaupt nicht mit seiner Auswahl, wurde eine **1** notiert. Durch diese Zuordnung war es möglich, für alle drei Fragebögen das arithmetische Mittel (siehe Kapitel 4.4.1) zu berechnen, um einen besseren Überblick zu bekommen und die Bögen zu vergleichen. Die Auswertung der Daten ergab für die drei Fragebögen die Ergebnisse, welche in Tabelle 5.2 dargestellt sind. Betrachtet man die Mittel der ersten

Auswertung	Daten	$\bar{x}$
Rebut	0, 0, 0, 0, 1, 0	0,167
Undercut	0, 0, 0, 0, 0, 1	0,167
Undermine	1, 1, 0, 0, 1, 0	0,5
Σ	13: 0 5: 1	0,278

Tabelle 5.2: Auswertungsdaten mit dem Verfahren aus Tabelle 5.1

beiden Fragebögen (bezüglich Rebut und Undercut), so gaben die Probanden sowohl mit ihrer eigenen Meinung, als auch mit der Auswahl ihrer Antwortmöglichkeit, überwiegend dasselbe Feedback. Beim dritten Fragebogen ist dies anders. Mit einem Mittel $\bar{x} = 0,5$ änderten die Hälfte der Probanden ihre Meinung bei der Auswahl einer der von D-BAS vorgeschlagenen Feedbackmöglichkeiten. Dies führte bei einem Probanden zur Auswahl der Antwort „Ich weiß es nicht. Zeige mir ein weiteres Argument“, andere Probanden wählten bei der Auswahl einer Möglichkeit eine durchweg verschiedene Aussage zur eigenen aus. Betrachtet wird beispielsweise die Aussage von Proband 017, dessen eigene Aussage „E-Autos sind für den Stadtverkehr empfehlenswert, da reichen 300 km auch aus. Nur für längere Fahrten könnte es knapp werden., nicht mit der gewählten Aussage „Ja, ich akzeptiere, dass es auch E-Autos gibt mit einer Reichweite von unter 150km.“ übereinstimmte. Seine Hauptaussage, dass E-Autos optimal für den Stadtverkehr sind und 300 Km ausreichen, wird in der gewählten Möglichkeit nicht aufgegriffen. Während der Durchführung des Testes beanstandeten viele der Probanden die sprachliche Umsetzung sowohl in den Chat-Blasen, als auch bei den Auswahlmöglichkeiten, welche von D-BAS vorgeschlagen wurden. Hierbei wurde kritisiert, dass die Formulierungen oftmals zu lang und unverständlich seien. Besonders fiel dies bei der Möglichkeit ins Gewicht, in der sowohl das Argument, als auch das Gegenargument akzeptiert wurden, jedoch eine stärkere Meinung zum eigenen Standpunkt bestand. Es ist zu beachten, dass der erste Test komplett ohne Colour-Coding (siehe Kapitel 3.5) durchgeführt wurde, weswegen es sehr wahrscheinlich zu den Verständnisproblemen im dritten Fragebogen gekommen sein könnte und die Hälfte der Probanden ihre Meinung bei der Auswahl geändert hat. Insgesamt betrachtet haben fünf der achtzehn Probanden bei der Auswahl einer Möglichkeit ihre Meinung geändert.

5.2 Auswertung des zweiten Tests

Bei der Auswertung des zweiten Tests wird das selbe Verfahren angewendet, wie zuvor beim ersten Test (siehe Tabelle 5.1). Dieses Mal haben pro Testbogen acht Probanden teilgenommen, sechs neue Probanden, die zuvor noch nicht teilgenommen haben, sowie zwei, die auch schon den ersten Testbogen durchlaufen sind. In Tabelle 5.3 sind die einzelnen Ergebnisse des zweiten Tests aufgeführt.

Auswertung	Probanden mit zweiter Teilnahme	Probanden mit erster Teilnahme	$\bar{x}$
Rebut	0, 0	0, 0, 0, 0, 0, 0	0
Undercut	0, 0	1, 1, 0, 1, 0, 0	0,375
Undermine	0, 0	0, 0, 0, 0, 0, 1	0,125
Σ	6: 0 0: 1	14: 0 4: 1	0,167

Tabelle 5.3: Auswertungsdaten zweiter Test

Erneut wurden die arithmetischen Mittel bestimmt. Vier der insgesamt vierundzwanzig Probanden änderten ihre Meinung bei der Auswahl einer der durch D-BAS vorgegebenen Antwortmöglichkeiten. Manche Probanden äußerten auch bei diesem Durchlauf, dass die Formulierungen der einzelnen Aussagen schwer verständlich und unübersichtlich seien. Außerdem wurde nachgefragt, ob die einzelnen farblichen Makierungen eine Bedeutung haben, was darauf schließen lässt, dass nicht direkt klar war, dass die Färbungen in den Auswahlmöglichkeiten in Bezug auf die einzelnen Aussagen in den Sprechblasen stehen. Des Weiteren wurde erneut geäußert, dass die grammatikalische Umsetzung der einzelnen Aussagen schwer verständlich sei.

5.3 Vergleich der Tests

Zum besseren Vergleich der beiden Testdurchläufe wurde den sechs Probanden, welche an beiden Testen teilgenommen haben, nach Beantwortung dieser zusätzlich ein Fragebogen vorgelegt. Dieser zusätzliche Fragebogen beinhaltete folgende Punkte, in denen der Proband die beiden Teste miteinander vergleichen sollte. Hierzu wurden zum zweiten Test folgende Fragen gestellt:

- Waren die Aussagen der Teilnehmer übersichtlicher?
- Halfen die farblichen Makierungen?
- Waren die Aussagen verständlicher formuliert?
- Hatten die Aussagen eine angenehme Länge?

Diese Punkte sollten von den Probanden erneut auf einer Likert-Skala bewertet werden, wobei diese erneut zwischen **1: trifft gar nicht zu** bis **5: trifft voll und ganz zu** abstimmen konnten. Außerdem wurde bezüglich Übersichtlichkeit und Verständlichkeit noch eine Vergleichsfrage gestellt, in der sich der Proband entweder für den ersten oder zweiten Test entscheiden sollte. Am Ende des Vergleichsbogens konnte der Proband noch seine eigene Meinung oder Bewertung äußern. Um diesen Zusatzbogen auszuwerten, wurde erneut das arithmetische Mittel der Likert-Skalen berechnet (siehe Tabelle 5.4, $n = 6$). Des Weiteren ist in Tabelle 5.5 der Vergleich der beiden Fragebögen dargestellt.

Mittel der Bewertungen des Vergleichstests	$\bar{x}$
Im zweiten Test waren die Aussagen der Teilnehmer übersichtlicher	3,33
Im zweiten Test waren die farblichen Markierungen hilfreich	3,5
Im zweiten Test waren die Aussagen verständlich formuliert	3,67
Im zweiten Test hatten die Aussagen eine angenehme Länge	4

Tabelle 5.4: Auswertungsdaten Vergleichsfragebogen

Der Proband sollte sich bei jeder Aussage für genau einen Bogen entscheiden (hierbei wird die die Auswahl des Probanden für den ersten Fragebogen mit einer 1 gezählt, für den zweiten Fragebogen mit einer 2). Die Daten in Tabelle 5.4 verdeutlichen, dass die Mehrheit der Probanden, welche beiden

Direkter Vergleich der beiden Fragebögen	Auswahl Test 1/ Test 2
Die Aussagen der Teilnehmer waren verständlicher	2, 2, 2, 2, 2, 2
Die Aussagen der Teilnehmer waren übersichtlicher	2, 2, 2, 2, 2, 1

Tabelle 5.5: Direkter Vergleich Bogen Nr. 1/2

Bögen getestet haben, die Änderungen im zweiten Fragebogen als Verbesserung zum ersten sehen. Diese unterstützen ebenfalls die Aussagen, welche die Probanden in ihrer eigenen Bewertung beschrieben haben. Es wurde geäußert, dass die Formulierungen besser geworden sind und die farbliche Gestaltung hilfreich sei. Außerdem seien die Aussagen kürzer und übersichtlicher. Als Kritikpunkt wurde hinzugefügt, dass die farbliche Gestaltung jedoch als Nachteil hat, dass Farben die Auswahl der Antwortmöglichkeit beeinflussen können. Daher sei die Auswahl der farblichen Gestaltung hier nur suboptimal. Es wurde darauf hingewiesen, dass die Farben gut lesbar sein sollen und auf helle Farben verzichtet werden soll. Bei anderen Probanden hingegen kam die farbliche Darstellung negativ an. Sie kritisierten, dass diese verwirrend sei. Durchweg zeigen die Daten aus Tabelle 5.5 aber, dass die Änderungen aus dem zweiten Test eine Verbesserung zu dem aus dem ersten Test darstellen. Die Mehrheit der Probanden sah diese als Verbesserung und lobte, dass „die Aussagen kürzer und übersichtlicher formuliert sind“ (Proband Vergleichsbogen 003). Als Hauptkritikpunkt wurde jedoch bei beiden Testdurchläufen die grammatikalische Formulierung von Aussagen und Auswahlmöglichkeiten geäußert. Kritisiert wurde dies als „komisch formuliertes Deutsch“ (Proband Vergleichsbogen 006). Um die beiden Teste bezüglich der Feedbackmöglichkeiten zu vergleichen, werden erneut die arithmetischen Mittel der Meinungsänderung von Probanden bezüglich der Auswahl einer von D-BAS präsentierten Möglichkeit gegenüber gestellt (siehe Tabelle 5.6). Vergleicht man die Werte, erkennt man, dass eine Verbesserung dieser hinsichtlich dem Einhalten der eigenen Meinung im zweiten Test gegeben ist.

Vergleich	Test Nr. 1 $\bar{x}$	Test Nr. 2 $\bar{x}$
Rebut	0,167	0
Undercut	0,167	0,375
Undermine	0,5	0,125

Tabelle 5.6: Vergleich Feedbackmöglichkeiten

5.4 Statistische Auswertung

Um die Daten statistisch auszuwerten, wird ein Konfidenzintervall für den Mittelwert $\bar{x}$ berechnet, um eine repräsentative Schätzung für eine größere Gruppe an Probanden zu ermitteln. Die Verteilung wird hierbei durch eine Bernoulli-Zufallsvariable beschrieben, da nur zwischen „Aussage und Auswahl stimmt überein: 0“ und „Aussage stimmt mit der Auswahl nicht überein: 1“ unterschieden wird [Lei04, S.128-129]. Angenommen ein Anteil p mit $0 \leq p \leq 1$ ändert seine Meinung während der Durchführung des Testes, dann bleibt ein Anteil von $(1 - p)$ Probanden, der seiner Meinung treu bleibt. Gesucht ist ein Intervall, in dem mit 95%-iger Sicherheit der Anteil p der Gruppe ihre Meinung ändert.

Zur Berechnung wird die Tschebyscheff-Ungleichung benutzt [Lei04, S.193].

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - p\right| > \varepsilon\right) \leq \frac{\text{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right)}{\varepsilon^2} = \frac{\text{Var}(X_i)}{n^2\varepsilon^2} \underset{\text{Bernoulli}}{X_i \text{ in}} \frac{(1-p) \cdot p \cdot n}{n^2\varepsilon^2} = \frac{1}{4n\varepsilon^2} \quad (5.1)$$

Nun wird das Konfidenzintervall mit Niveau 95% für den zweiten Test bestimmt, indem $\alpha = 0,95$, $\bar{x} = \frac{1}{n}\sum_{i=1}^{24} X_i = 0,167$ (siehe Tabelle 5.3) und $n = 24$ gewählt wird, wobei n die Anzahl an teilgenommenen Probanden ist. Es wird gemäß (5.1)

$$\alpha = \frac{1}{4n\varepsilon^2} \quad (5.2)$$

gesetzt, wodurch folgendes erhalten wird:

$$\begin{aligned} 0,95 &= \frac{1}{4 \cdot 24 \cdot \varepsilon^2} \\ \varepsilon &= \sqrt{\frac{1}{0,95 \cdot 4 \cdot 24}} \\ \varepsilon &\approx 0,1047 \end{aligned} \quad (5.3)$$

Für das Konfidenzintervall gilt $[\frac{\sum_{i=1}^n X_i}{n} - \varepsilon ; \frac{\sum_{i=1}^n X_i}{n} + \varepsilon]$ [Kä10, S.263], was nach Einsetzen der vorhandenen Werte folgendes Konfidenzintervall für das gesuchte p ergibt:

$$p \in [0,023; 0,271].$$

Die Auswertung zeigt, dass mit 95%-iger Sicherheit p , der globale Anteil an D-BAS Nutzern, die ihre Meinung ändern, nicht über 0,271 steigt. Auf gleiche Weise wird ebenfalls zum besseren Vergleich das Konfidenzintervall für den ersten Test berechnet. Dies ergibt

$$p_1 \in [0,157; 0,399] \text{ und einem Mittelwert } \bar{x}_1 = 0,278.$$

Wie zu erkennen ist, liegt die obere Intervallsgrenze beim ersten Test höher als beim zweiten Test. Daraus lässt sich schlussfolgern, dass Verbesserungen wie dem Colour-Coding oder der Verkürzung der Antwortmöglichkeiten positiv messbare Besserungen für den Nutzer von D-BAS sind und sich der Anteil der Meinungsänderungen verringert hat.

5.5 Verbesserungen

Um in Zukunft D-BAS benutzerfreundlicher zu gestalten, empfiehlt sich die Überarbeitung der Formulierungen innerhalb des Systems. Hierbei sollten Satzbau und Wortwahl bei Angriffen des Systems geändert werden, damit die Verknüpfung der einzelnen Aussagen leichter für den Nutzer zu verstehen sind. Um dies zu verdeutlichen, wird ein Verbesserungsvorschlag für jeden der drei Angriffsfälle anhand eines Beispiels präsentiert.

Rebut Betrachtet wird nun folgende Aussage:

„Andere Teilnehmer stimmen zu, dass die Reichweite von Elektroautos ausreichend für mindestens 300 Km ist, aber Sie haben eine stärkere Aussage zur Ablehnung davon, dass E-Autos optimal für den Stadtverkehr sind. Sie sagen, dass die Ladezeit der Batterie bis zu 12 h dauern kann und so lange man tagsüber nicht warten kann“.

Viele der Probanden waren mit der Negation „stärkere Aussage zur Ablehnung“ überfordert oder brauchten einige Zeit, um diese zu verstehen. Eventuell könnte hierbei folgende Formulierung benutzt werden:

„Andere Teilnehmer stimmen zu, dass die Reichweite von Elektroautos ausreichend für mindestens 300 Km ist, jedoch denken sie nicht, dass E-Autos optimal für den Stadtverkehr sind. Sie sagen, dass die Ladezeit der Batterien bis zu 12 h dauern kann und so lange man tagsüber nicht warten kann“.

Hilfreich für den Nutzer wären außerdem die Hervorhebung einzelner Signalwörter, welche die Aussage einer Verknüpfung oder eines Standpunktes verdeutlichen. Dies könnte beispielsweise durch fettgedruckte Wörter umgesetzt werden. Das gerade genannte Beispiel würde sich wie folgt verändern:

„Andere Teilnehmer **stimmen zu**, dass die Reichweite von Elektroautos ausreichend für mindestens 300 Km ist, jedoch denken sie **nicht**, dass E-Autos optimal für den Stadtverkehr sind. Sie sagen, dass die Ladezeit der Batterien bis zu 12 h dauern kann und so lange man tagsüber nicht warten kann“.

Undercut Beim Undercut sieht dies ein wenig anders aus. Die Verknüpfung der einzelnen Aussagen waren in diesem Fall für den Nutzer weitestgehend verständlich. Betrachtet wird das Beispiel:

„Andere Teilnehmer stimmen zu, dass die Reichweite von Elektroautos ausreichend für mindestens 300 Km ist, aber sie glauben, dass es kein gutes Argument dafür ist, dass E-Autos optimal für den Stadtverkehr sind. Sie denken, dass in der Stadt Fahrräder und öffentliche Verkehrsmittel besser sind“.

Jedoch würde es sich erneut anbieten, Signalwörter hervorzuheben.

„Andere Teilnehmer **stimmen zu**, dass die Reichweite von Elektroautos ausreichend für mindestens 300 Km ist, aber sie glauben, dass es **kein gutes Argument** dafür ist, dass E-Autos optimal für den Stadtverkehr sind. Sie **denken**, dass in der Stadt Fahrräder und öffentliche Verkehrsmittel besser sind“.

Undermine Beim Undermine traten hingegen erneut Verständnisprobleme bei den Probanden auf. Zusätzlich wurde der Satzbau kritisiert. Betrachtet wird das Beispiel:

„Andere Teilnehmer denken, dass E-Autos eine Reichweite von mindestens 300 Km haben ist keine gute Idee, weil nicht klar ist, wie lange die Batterieeinheiten der jeweiligen E-Autos halten“.

Durch Änderungen des Satzbaus und Hervorheben der Signalwörter folgt folgende Verbesserung:

„Andere Teilnehmer denken **nicht**, dass E-Autos eine Reichweite von mindestens 300 Km haben, weil nicht klar ist, wie lange die Batterieeinheiten der jeweiligen E-Autos halten“.

Antwortmöglichkeiten Bei den vorgegebenen Antwortmöglichkeiten wurde nicht immer verstanden, dass „deren Standpunkt“ die Ablehnung des eigenen Standpunktes beschreibt. Betrachtet wird hierzu die Antwortmöglichkeit:

„Ja, ich akzeptiere deren Argument und ich akzeptiere, dass es ein Argument für deren Standpunkt ist. Jedoch habe ich ein viel stärkeres Argument für meinen Standpunkt.“

Die Phrase „deren Standpunkt“ sollte überdacht werden und eventuell durch die Angabe „gegen meinen Standpunkt“ ersetzt werden. Die oben genannte Antwortmöglichkeit würde sich wie folgt verändern

„Ja, ich akzeptiere deren Argument und ich akzeptiere, dass es ein Argument **gegen** meinen Standpunkt ist. Jedoch habe ich ein viel stärkeres Argument für meinen Standpunkt.“

Diese Änderung würde außerdem bewirken, dass nur noch zwei verschiedene Farben für das Colour-Coding verwendet werden müssen (für „meinen Standpunkt“ und „deren Argument“). Die Verbesserungen für die Antwortmöglichkeiten könnten wie folgt verändert werden:

- Nein, ich lehne deren Argument ab.
- Ja, ich akzeptiere deren Argument.
- Ja, ich akzeptiere deren Argument, aber ich glaube nicht, dass es ein gutes Argument gegen meinen Standpunkt ist.
- Ja, ich akzeptiere deren Argument und ich akzeptiere, dass es ein gutes Argument gegen meinen Standpunkt ist, jedoch habe ich ein stärkeres Argument für meinen Standpunkt.
- Ich weiß es nicht. Zeige mir ein weiteres Argument.

Um die Änderungen zu verdeutlichen, wird das Beispiel in Abbildung 5.1 betrachtet. Beim Colour-Coding muss überlegt werden, ob die gewählten Farben den Nutzer nicht bei seiner Entscheidung beeinflusst. Da der Mehrheit der Probanden diese Änderung beim Verständnis jedoch geholfen hat, ist die generelle Idee dahinter vorteilhaft. Außerdem wird beim Ersetzen der einzelnen Aussagen durch Phrasen wie „mein Standpunkt“ sowie „deren Argument“ die Satzlänge der einzelnen Auswahlmöglichkeiten, die vom System vorgegeben werden, erheblich verkürzt, was dem Nutzer positiv auffiel und zum Verständnis beigetragen hat (siehe Ergebnisse Vergleichsbogen Tabelle 5.5). Da D-BAS für jeden Nutzer verständlich und leicht zu bedienen sein soll, darf die gewählte Wort- und Ausdruckswahl

nicht zu kompliziert und fachspezifisch sein. Zusammengefasst sollten folgende Punkte überdacht und daraufhin eventuell überarbeitet werden:

- Formulierungen, die D-BAS benutzt, um verschiedene Aussagen zu verknüpfen
- Vermeidung von komplizierten Negationen
- Farbauswahl beim Colour-Coding
- Verdeutlichung von Akzeptieren oder Ablehnung durch entsprechend, hervorgehobene Signalwörter
- Ersetzung von „deren Standpunkt“ durch „gegen meinen Standpunkt“

Ich bin der Meinung, dass E-Autos optimal für den Stadtverkehr sind
16 weitere Teilnehmer/innen mit dieser Meinung.

Jetzt

Ich akzeptiere, dass E-Autos optimal für den Stadtverkehr sind weil E-Autos eine Reichweite von mindestens 300 Km haben
25 weitere Teilnehmer/innen mit dieser Meinung.

Andere Teilnehmer **stimmen zu**, dass E-Autos eine Reichweite von mindestens 300 Km haben, jedoch denken sie **nicht**, dass E-Autos optimal für den Stadtverkehr sind. Sie sagen, dass die Ladezeit der Batterie bis zu 12 Stunden dauern kann und man insbesondere tagsüber diese Zeit oftmals nicht hat.

Was denken Sie darüber?

☐ Nein, ich lehne deren Argument ab .

☐ Ja, ich akzeptiere deren Argument.

☐ Ja, ich akzeptiere deren Argument, aber ich glaube nicht, dass es ein gutes Argument gegen meinen Standpunkt ist.

☐ Ja, ich akzeptiere deren Argument und ich akzeptiere, dass es ein Argument gegen meinen Standpunkt ist. Jedoch habe ich ein viel stärkeres Argument für meinen Standpunkt.

☐ Ich weiß es nicht. Zeige mir ein weiteres Argument.

Abbildung 5.1: Beispiel mit möglichen Verbesserungen

Kapitel 6

Fazit und Ausblick

6.1 Fazit

Diese Bachelorthesis befasste sich mit der Analyse der Feedbackmöglichkeiten des dialog-basierten Argumentationssystems D-BAS. Hierbei sollte herausgefunden werden, ob ein Nutzer des Systems ohne dessen Vorgaben in deiner Diskussion genau so reagiert, als wenn er Antwortmöglichkeiten zur Interaktion vorgegeben bekommt. Um dieser Fragestellung nachzugehen, wurden zwei Testrunden durchgeführt, in denen Probanden sowohl mit als auch ohne Vorgaben des Systems auf eine angefangene Diskussion reagieren sollten. Nach den Testdurchläufen wurde diese ausgewertet und Verbesserungsvorschläge für D-BAS entwickelt. Durch Änderungen, wie dem Colour-Coding sowie der Verkürzung des Satzbau durch die Ersetzung der einzelnen Aussagen durch kurze Phrasen („deren Argument“, „mein Standpunkt“), wird dem Benutzer eine gute Möglichkeit zur Interaktion mit dem System geboten. Jedoch ist aufgefallen, dass noch einige Probleme im Verständnisbereich auftreten. Hierzu gehört als wichtigster Punkt die Formulierungen der Verknüpfungen zwischen einzelnen Aussagen sowie der lange Satzbau der Antwortmöglichkeiten. Die Testläufe haben gezeigt, dass der durchschnittliche Internetnutzer einige Probleme mit der aktuellen Version des Prototypen hat und seine Meinung oftmals noch nicht auf die Auswahlmöglichkeiten des Systems übertragen kann, was zu falschen Eingaben der eigenen Meinung während einer Diskussion führen kann. Das Problem ist hierbei allerdings nicht die Möglichkeiten des Feedbacks durch den Benutzer, sondern die fachspezifischen Formulierungen innerhalb des Systems. Dargelegt wurden Verbesserungen im Bereich der Formulierungen bezüglich Verknüpfung von Aussagen, welche diese leichter verständlich machen sollen. Außerdem wurde die Einführung von fett gedruckten Signalwörtern vorgeschlagen, die das Ablehnen oder Akzeptieren unterstreichen und somit die Haltung des Nutzer für ihn direkt ersichtlich machen sollen. Des Weiteren wurde die Änderung der Phrase „deren Standpunkt“ zu „gegen meinen Standpunkt“ angeregt, die dem Nutzer vorteilhafter die verschiedenen Standpunkte verdeutlichen soll. Dies trägt zur Optimierung des Colour-Codings bei. Die Anzahl der Farben, welche benötigt

werden, wird hierbei reduziert, was zur Verbesserung der Übersichtlichkeit für den Nutzer führen könnte. Letztendlich wurde festgestellt, dass D-BAS in diesen genannten Punkten noch Möglichkeiten zur Überarbeitung besitzt, welche in Zukunft überdacht und verbessert werden können, damit es dem Nutzer einfacher fällt, in einer mit D-BAS geführten Diskussion zu interagieren.

6.2 Ausblick

In Zukunft wird es weitere Veränderungen am Prototypen von D-BAS geben, welche daraufhin erneut getestet werden müssen. Das Grundgerüst erlaubt dem Benutzer, mit dem System durch vorgegebene Auswahlmöglichkeiten zu interagieren und somit eine Online-Diskussion zu führen. Jedoch sollten noch die geschilderten Kritikpunkte überarbeitet werden, um das System weiter zu optimieren. Um ein repräsentativeres Testergebnis zu bekommen, sollte in Zukunft ein Usability-Test mit mehreren Nutzern durchgeführt werden, um weitere Probleme aufzudecken, die der Prototyp eventuell noch beinhaltet und zu überprüfen, ob die in dieser Arbeit geschilderten Vorschläge eine Verbesserung bezüglich Verständnis und Interaktion für den Nutzer darstellen. Das Ziel sollte sein, D-BAS dahingehend zu verändern, dass der Nutzer leicht und verständlich die Möglichkeit hat, seine eigene Meinung zu einem Thema durch die Auswahl einer vorgegebenen Möglichkeit zu äußern und somit eine Diskussion mit anderen Nutzern zu führen, unabhängig von Zeit und Ort.

Anhang

(1) Fragebogen Pre-Test Rebut (Fragebogen Undercut und Undermine auf der CD)

(2) Fragebogen erster Test Rebut (Fragebogen Undercut und Undermine auf der CD)

(3) Fragebogen zweiter Test Rebut (Fragebogen Undercut und Undermine auf der CD)

(4) Fragebogen zur Bewertung zwischen dem ersten und zweiten Test



D-BAS

Dialog-Basiertes Argumentations System

D-BAS - Dialog-Basiertes Argumentations System (Studie im Rahmen der Bachelorarbeit) PRETEST

Studie im Rahmen der Bachelorarbeit: "Dialog-Basiertes Argumentations System" (D-BAS)

Liebe Teilnehmerin, lieber Teilnehmer,

dies ist eine Studie im Rahmen meiner Bachelorarbeit mit dem Thema "Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems".

In dieser Studie wird Dir der Prototyp einer dialog-basierten Argumentationssoftware vorgestellt (D-BAS).

Lies Dir bitte alle Fragen in Ruhe durch und makiere jeweils deine gewählte Antwort.

In dieser Studie sollst du deine Gedanken bei der Beantwortung der Fragen laut und offen aussprechen.

Bitte beachte, dass dies KEIN Test ist. Du wirst nicht geprüft.

Für Fragen stehe ich Dir jeder Zeit zur Verfügung.

Nun wünsche ich Dir viel Spaß bei der Umfrage.

Datenschutz

Einverständniserklärung zur Mitwirkung an der Studie zu meiner Bachelorarbeit
"Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems" von

Daniel Braune
daniel.braune@hhu.de
Institut für Informatik

Erklärung zum Datenschutz:

Vielen Dank, dass Du Dir die Zeit nimmst, an meiner Studie teilzunehmen. Die Studie dient rein wissenschaftlichen Zielen. Die erhobenen Daten werden nur in anonymisierter Form ausgewertet und die Ergebnisse werden ebenfalls nur in anonymisierter Form veröffentlicht. Alle erhobenen Daten werden vertraulich behandelt und nicht zu Werbezwecken verwendet. Zu keiner Zeit erfolgt eine Weitergabe der erhobenen Daten an unbefugte Dritte.

Ich stimme zu, dass meine persönlichen Daten nur in anonymisierter Form an außen stehende Institutionen zu Forschungszwecken weitergegeben werden.

Mir ist bekannt, dass während der Teilnahme Daten erhoben werden wobei es sich konkret um die Aufzeichnung von Protokolldaten der verwendeten Software, von mir selbst produzierten Inhalten und Bildschirminteraktionen sowie Daten aus einer abschließenden Befragung handelt. Des weiteren ist mir bekannt, dass die erhobenen Daten im Rahmen einer Abschlussarbeit in anonymisierter Form ausgewertet und nur für wissenschaftliche Zwecke verwendet werden.

Ich hatte ausreichend Zeit, mich zur Teilnahme an der Studie zu entscheiden und weiß, dass die Teilnahme freiwillig ist. Ich weiß, dass ich jederzeit und ohne Angabe von Gründen diese Zustimmung widerrufen kann, ohne dass dadurch Nachteile für mich entstehen.

Einverständnis der Datenschutzerklärung *

☐ Ich habe die Datenschutzerklärung gelesen und stimme dieser zu.

Start Testfragebogen

Nun startet der Testfragebogen.

Bitte beantworte hierzu die folgenden Fragen.

Nach Abschluss des Testfragebogens wirst du benachrichtigt.

Geschlecht

Ich bin *

☐ männlich

☐ weiblich

Alter

Wie alt bist du? *

☐ 18-21

☐ 22-25

☐ 26-29

☐ 30-34

☐ 35-39

☐ 40+

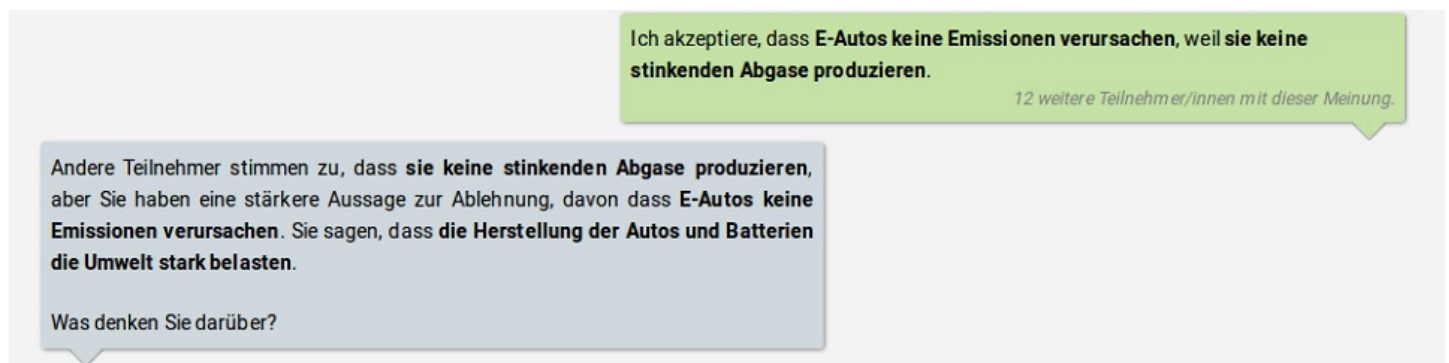
Beruf

Welchen höchsten allgemeinbildenden Schulabschluss hast du? *

- ☐ bin noch Schüler(in)
- ☐ Schule beendet ohne Abschluss
- ☐ Hauptschulabschluss, Volksschulabschluss
- ☐ Realschulabschluss, Mittlere Reife, Fachschulreife oder Polytechnikum 10. Klasse
- ☐ Fachhochschulreife (Abschluss einer Fachoberschule, etc.)
- ☐ Abitur oder erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)
- ☐ Hochschulabschluss (Universität oder Fachhochschule)
- ☐ anderes

Abbildung 1: D-BAS

Abbildung 1



Was denkst du darüber? *

Abbildung 1: D-BAS

Abbildung 1

Ich akzeptiere, dass **E-Autos keine Emissionen verursachen**, weil **sie keine stinkenden Abgase produzieren**.
12 weitere Teilnehmer/innen mit dieser Meinung.

Andere Teilnehmer stimmen zu, dass **sie keine stinkenden Abgase produzieren**, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass **E-Autos keine Emissionen verursachen**. Sie sagen, dass **die Herstellung der Autos und Batterien die Umwelt stark belasten**.

Was denken Sie darüber?

Für welche dieser Antwortmöglichkeiten würdest du dich entscheiden? *

- ☐ Nein, ich lehne ab, dass die Herstellung der Autos und Batterien die Umwelt stark belasten.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten, aber ich glaube nicht, dass es ein gutes Argument dagegen ist, dass E-Autos keine Emissionen verursachen.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten und ich akzeptiere, dass es ein Argument dagegen ist, dass E-Autos keine Emissionen verursachen. Jedoch habe ich ein viel stärkeres Argument dafür, dass E-Autos keine Emissionen verursachen.
- ☐ Ich weiß es nicht. Zeige mir ein weiteres Argument.
- ☐ Keine der angegebenen Antwortmöglichkeiten, sondern:

Testfragebogen beendet

Nun sollst du den Fragebogen, welchen du gerade beantwortet hast, bewerten. Klicke dich hierzu durch die Bewertungsfragen.

Bewertung

Bewerte bitte folgende Aussagen. *

Bewerte auf einer Skala von 1: "stimme überhaupt nicht zu" bis 5: "stimme voll und ganz zu"

	1	2	3	4	5
Die Grafiken waren verständlich.	☐	☐	☐	☐	☐
Die Frageseiten waren übersichtlich.	☐	☐	☐	☐	☐
Ich hatte keine Verständnisprobleme.	☐	☐	☐	☐	☐

Solltest du mindestens einmal 1,2 oder 3 gewählt haben, warum hast du diese Bewertung gewählt? Bitte schildere kurz deine Kritik.

Anmerkungen

Hast Du weitere Anmerkungen?

Vielen Dank!

Vielen Dank für deine Teilnahme an dieser Studie.

Du hast mir dadurch beim Anfertigen meiner Bachelorarbeit sehr geholfen.

Daniel Braune

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



D-BAS

Dialog-Basiertes Argumentations System

D-BAS - Dialog-Basiertes Argumentations System (Studie im Rahmen der Bachelorarbeit)

Studie im Rahmen der Bachelorarbeit: "Dialog-Basiertes Argumentations System" (D-BAS)

Liebe Teilnehmerin, lieber Teilnehmer,

dies ist eine Studie im Rahmen meiner Bachelorarbeit mit dem Thema "Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems".

In dieser Studie wird Dir der Prototyp einer dialog-basierten Argumentationssoftware vorgestellt (D-BAS).

Lies Dir bitte alle Fragen in Ruhe durch und makiere jeweils deine gewählte Antwort.

In dieser Studie sollst du deine Gedanken bei der Beantwortung der Fragen laut und offen aussprechen.

Bitte beachte, dass dies KEIN Test ist. Du wirst nicht geprüft.

Für Fragen stehe ich Dir jeder Zeit zur Verfügung.

Nun wünsche ich Dir viel Spaß bei der Umfrage.

Datenschutz

Einverständniserklärung zur Mitwirkung an der Studie zu meiner Bachelorarbeit
"Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems" von

Daniel Braune
daniel.braune@hhu.de
Institut für Informatik

Erklärung zum Datenschutz:

Vielen Dank, dass Du Dir die Zeit nimmst, an meiner Studie teilzunehmen. Die Studie dient rein wissenschaftlichen Zielen. Die erhobenen Daten werden nur in anonymisierter Form ausgewertet und die Ergebnisse werden ebenfalls nur in anonymisierter Form veröffentlicht. Alle erhobenen Daten werden vertraulich behandelt und nicht zu Werbezwecken verwendet.

Zu keiner Zeit erfolgt eine Weitergabe der erhobenen Daten an unbefugte Dritte.

Ich stimme zu, dass meine persönlichen Daten nur in anonymisierter Form an außen stehende Institutionen zu Forschungszwecken weitergegeben werden.

Mir ist bekannt, dass während der Teilnahme Daten erhoben werden wobei es sich konkret um die Aufzeichnung von Protokolldaten der verwendeten Software, von mir selbst produzierten Inhalten und Bildschirminteraktionen sowie Daten aus einer abschließenden Befragung handelt.

Des weiteren ist mir bekannt, dass die erhobenen Daten im Rahmen einer Abschlussarbeit in anonymisierter Form ausgewertet und nur für wissenschaftliche Zwecke verwendet werden.

Ich hatte ausreichend Zeit, mich zur Teilnahme an der Studie zu entscheiden und weiß, dass die Teilnahme freiwillig ist. Ich weiß, dass ich jederzeit und ohne Angabe von Gründen diese Zustimmung widerrufen kann, ohne dass dadurch Nachteile für mich entstehen.

Einverständnis der Datenschutzerklärung *

☐

Ich habe die Datenschutzerklärung gelesen und stimme dieser zu.

Geschlecht

Ich bin *

☐

männlich

☐

weiblich

Alter

Wie alt bist du? *

- ☐ 18-21
- ☐ 22-25
- ☐ 26-29
- ☐ 30-34
- ☐ 35-39
- ☐ 40+

Beruf

Welchen höchsten allgemeinbildenden Schulabschluss hast du? *

- ☐ bin noch Schüler(in)
- ☐ Schule beendet ohne Abschluss
- ☐ Hauptschulabschluss, Volksschulabschluss
- ☐ Realschulabschluss, Mittlere Reife, Fachschulreife oder Polytechnikum 10. Klasse
- ☐ Fachhochschulreife (Abschluss einer Fachoberschule, etc.)
- ☐ Abitur oder erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)
- ☐ Hochschulabschluss (Universität oder Fachhochschule)
- ☐ anderes

Diskussionsbeginn

Auf den folgenden zwei Seiten findest du eine angefangene Diskussion über Elektroautos (E-Autos), welche mit dem dialog-basierten Argumentationssystem D-BAS geführt wurde.

Lies dir hierzu die bisherige Konversation durch und beantworte die jeweilige Fragestellung.

Abbildung 1: D-BAS

Abbildung 1

The screenshot shows a survey interface with two opinion bubbles. The top bubble is green and contains the text: "Ich akzeptiere, dass **E-Autos keine Emissionen verursachen**, weil sie **keine stinkenden Abgase produzieren**." Below this text, in smaller font, it says "12 weitere Teilnehmer/innen mit dieser Meinung." The bottom bubble is light blue and contains the text: "Andere Teilnehmer stimmen zu, dass **sie keine stinkenden Abgase produzieren**, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass **E-Autos keine Emissionen verursachen**. Sie sagen, dass **die Herstellung der Autos und Batterien die Umwelt stark belasten**." Below this text, it says "Was denken Sie darüber?"

Ich akzeptiere, dass **E-Autos keine Emissionen verursachen**, weil sie **keine stinkenden Abgase produzieren**.
12 weitere Teilnehmer/innen mit dieser Meinung.

Andere Teilnehmer stimmen zu, dass **sie keine stinkenden Abgase produzieren**, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass **E-Autos keine Emissionen verursachen**. Sie sagen, dass **die Herstellung der Autos und Batterien die Umwelt stark belasten**.
Was denken Sie darüber?

Was ist deine Meinung zur Aussage der anderen Teilnehmer? *

Abbildung 1: D-BAS

Abbildung 1

The screenshot shows a survey interface with two opinion bubbles. The top bubble is green and contains the text: "Ich akzeptiere, dass **E-Autos keine Emissionen verursachen**, weil sie **keine stinkenden Abgase produzieren**." Below this text, in smaller font, it says "12 weitere Teilnehmer/innen mit dieser Meinung." The bottom bubble is light blue and contains the text: "Andere Teilnehmer stimmen zu, dass **sie keine stinkenden Abgase produzieren**, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass **E-Autos keine Emissionen verursachen**. Sie sagen, dass **die Herstellung der Autos und Batterien die Umwelt stark belasten**." Below this text, it says "Was denken Sie darüber?"

Ich akzeptiere, dass **E-Autos keine Emissionen verursachen**, weil sie **keine stinkenden Abgase produzieren**.
12 weitere Teilnehmer/innen mit dieser Meinung.

Andere Teilnehmer stimmen zu, dass **sie keine stinkenden Abgase produzieren**, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass **E-Autos keine Emissionen verursachen**. Sie sagen, dass **die Herstellung der Autos und Batterien die Umwelt stark belasten**.
Was denken Sie darüber?

Für welche dieser Antwortmöglichkeiten würdest du dich entscheiden? *

- ☐ Nein, ich lehne ab, dass die Herstellung der Autos und Batterien die Umwelt stark belasten.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten, aber ich glaube nicht, dass es ein gutes Argument dagegen ist, dass E-Autos keine Emissionen verursachen.
- ☐ Ja, ich akzeptiere, dass die Herstellung der Autos und Batterien die Umwelt stark belasten und ich akzeptiere, dass es ein Argument dagegen ist, dass E-Autos keine Emissionen verursachen. Jedoch habe ich ein viel stärkeres Argument dafür, dass E-Autos keine Emissionen verursachen.
- ☐ Ich weiß es nicht. Zeige mir ein weiteres Argument.
- ☐ Keine der angegebenen Antwortmöglichkeiten, sondern:

Vielen Dank!

Vielen Dank für deine Teilnahme an dieser Studie.

Du hast mir dadurch beim Anfertigen meiner Bachelorarbeit sehr geholfen.

Daniel Braune

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



D-BAS

Dialog-Basiertes Argumentations System

D-BAS - Dialog-Basiertes Argumentations System (Studie im Rahmen der Bachelorarbeit)

Studie im Rahmen der Bachelorarbeit: "Dialog-Basiertes Argumentations System" (D-BAS)

Liebe Teilnehmerin, lieber Teilnehmer,

dies ist eine Studie im Rahmen meiner Bachelorarbeit mit dem Thema "Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems".

In dieser Studie wird Dir der Prototyp einer dialog-basierten Argumentationssoftware vorgestellt (D-BAS).

Lies Dir bitte alle Fragen in Ruhe durch und makiere jeweils deine gewählte Antwort.

In dieser Studie sollst du deine Gedanken bei der Beantwortung der Fragen laut und offen aussprechen.

Bitte beachte, dass dies KEIN Test ist. Du wirst nicht geprüft.

Für Fragen stehe ich Dir jeder Zeit zur Verfügung.

Nun wünsche ich Dir viel Spaß bei der Umfrage.

Datenschutz

Einverständniserklärung zur Mitwirkung an der Studie zu meiner Bachelorarbeit
"Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems" von

Daniel Braune
daniel.braune@hhu.de
Institut für Informatik

Erklärung zum Datenschutz:

Vielen Dank, dass Du Dir die Zeit nimmst, an meiner Studie teilzunehmen. Die Studie dient rein wissenschaftlichen Zielen. Die erhobenen Daten werden nur in anonymisierter Form ausgewertet und die Ergebnisse werden ebenfalls nur in anonymisierter Form veröffentlicht. Alle erhobenen Daten werden vertraulich behandelt und nicht zu Werbezwecken verwendet. Zu keiner Zeit erfolgt eine Weitergabe der erhobenen Daten an unbefugte Dritte.

Ich stimme zu, dass meine persönlichen Daten nur in anonymisierter Form an außen stehende Institutionen zu Forschungszwecken weitergegeben werden.

Mir ist bekannt, dass während der Teilnahme Daten erhoben werden wobei es sich konkret um die Aufzeichnung von Protokolldaten der verwendeten Software, von mir selbst produzierten Inhalten und Bildschirminteraktionen sowie Daten aus einer abschließenden Befragung handelt. Des weiteren ist mir bekannt, dass die erhobenen Daten im Rahmen einer Abschlussarbeit in anonymisierter Form ausgewertet und nur für wissenschaftliche Zwecke verwendet werden.

Ich hatte ausreichend Zeit, mich zur Teilnahme an der Studie zu entscheiden und weiß, dass die Teilnahme freiwillig ist. Ich weiß, dass ich jederzeit und ohne Angabe von Gründen diese Zustimmung widerrufen kann, ohne dass dadurch Nachteile für mich entstehen.

Einverständnis der Datenschutzerklärung *

☐ Ich habe die Datenschutzerklärung gelesen und stimme dieser zu.

Geschlecht

Ich bin *

☐ männlich

☐ weiblich

Alter

Wie alt bist du? *

18-21

☐

22-25

☐

26-29

☐

30-34

☐

35-39

☐

40+

☐

Beruf

Welchen höchsten allgemeinbildenden Schulabschluss hast du? *

☐

bin noch Schüler(in)

☐

Schule beendet ohne Abschluss

☐

Hauptschulabschluss, Volksschulabschluss

☐

Realschulabschluss, Mittlere Reife, Fachschulreife oder Polytechnikum 10. Klasse

☐

Fachhochschulreife (Abschluss einer Fachoberschule, etc.)

☐

Abitur oder erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)

☐

Hochschulabschluss (Universität oder Fachhochschule)

☐

anderes

Diskussionsbeginn

Auf den folgenden zwei Seiten findest du eine angefangene Diskussion über Elektroautos (E-Autos), welche mit dem dialog-basierten Argumentationssystem D-BAS geführt wurde.

Lies dir hierzu die bisherige Konversation durch und beantworte die jeweilige Fragestellung.

Abbildung 1: D-BAS

Abbildung 1

The screenshot shows a simulated discussion interface. At the top, a green speech bubble contains the text: "Ich akzeptiere, dass E-Autos optimal für den Stadtverkehr sind weil E-Autos eine Reichweite von mindestens 300 Km haben". Below this, in smaller text, it says "31 weitere Teilnehmer/innen mit dieser Meinung." Below the green bubble, a grey speech bubble contains the text: "Andere Teilnehmer stimmen zu, dass E-Autos eine Reichweite von mindestens 300 Km haben, aber Sie haben eine stärkere Aussage zur Ablehnung, davon dass E-Autos optimal für den Stadtverkehr sind. Sie sagen, dass die Ladezeit der Batterie bis zu 12 Stunden dauern kann und man insbesondere tagsüber diese Zeit oftmals nicht hat." Below the grey bubble, it says "Was denken Sie darüber?"

Was ist deine Meinung zur Aussage der anderen Teilnehmer? *

Abbildung 1: D-BAS

Abbildung 1

The screenshot shows the same discussion interface as above, but with a response form below the grey speech bubble. The form contains five radio button options:

- ☐ Nein, ich lehne deren Argument ab.
- ☐ Ja, ich akzeptiere deren Argument.
- ☐ Ja, ich akzeptiere deren Argument, aber ich glaube nicht, dass es ein gutes Argument für deren Standpunkt ist.
- ☐ Ja, ich akzeptiere deren Argument und ich akzeptiere, dass es ein Argument für deren Standpunkt ist. Jedoch habe ich ein viel stärkeres Argument für meinen Standpunkt.
- ☐ Ich weiß es nicht. Zeige mir ein weiteres Argument.

Für welche dieser Antwortmöglichkeiten würdest du dich entscheiden? *

☐ Antwort 1

☐ Antwort 2

☐ Antwort 3

☐ Antwort 4

☐ Antwort 5

Vielen Dank!

Vielen Dank für deine Teilnahme an dieser Studie.

Du hast mir dadurch beim Anfertigen meiner Bachelorarbeit sehr geholfen.

Daniel Braune

» Umleitung auf Schlussseite von Umfrage Online

Bewertungsbogen: "Dialog-Basiertes Argumentations System" (D-BAS)

Lieber Teilnehmer,

du hast bereits zweimal an meiner Studie im Rahmen meiner Bachelorarbeit

Analyse der Benutzerinteraktionen eines dialog-basierten Argumentationssystems

teilgenommen.

Nun sollst du die Aussagen im gerade bearbeiteten Umfragebogen (hier **Fragebogen Nr. 2** genannt) kurz bewerten.

Fülle dazu bitte den nachfolgenden Bewertungsbogen aus.

Für Fragen und Anmerkungen stehe ich dir jederzeit zur Verfügung.

Umfragebogen Nr.:

Daniel Braune

Institut für Informatik

Heinrich-Heine-Universität Düsseldorf

1 Alter

Wie alt bist du?

Ich bin _____ Jahre.

2 Geschlecht

Ich bin ☐ männlich ☐ weiblich

3 Bewerte nun die folgenden Aussagen durch Ankreuzen.

Beziehe dich bitte nur auf **Fragebogen Nr. 2**

Hierbei ist 1: **trifft überhaupt nicht zu** und 5: **trifft voll und ganz zu**

	1	2	3	4	5
Der Aussagen der Teilnehmer sind insgesamt übersichtlich					
Die farblichen Markierungen helfen beim Verständnis					
Die Aussagen der einzelnen Teilnehmer sind verständlicher formuliert					
Die Aussagen der Teilnehmer haben eine angenehme Länge					

4 Gegenüberstellung der beiden Fragebögen

Entscheide dich nun, welche Aussage auf welchen Fragebogen zutrifft. Kreuze hierzu an.

	Fragebogen Nr. 1	Fragebogen Nr. 2
Die Aussagen der Teilnehmer waren verständlicher		
Diese Aussagen der Teilnehmer waren übersichtlicher		

5 Bewertung

Sind deiner Meinung nach die Darstellungen und Formulierungen der einzelnen Aussagen der Teilnehmer in **Fragebogen Nr. 2** eine Verbesserung zu denen in **Fragebogen Nr. 1**?

☐ **Ja**, denn _____

☐ **Nein**, denn _____

6 Anmerkungen

Solltest du weitere Anmerkungen haben, kannst du diese hier aufschreiben.

Vielen Dank für deine Teilnahme!

Literaturverzeichnis

- [Bay07] BAYER, Klaus: *Argument und Argumentation: Logische Grundlagen der Argumentationsanalyse*. 2007. Vandenhoeck + Ruprecht Verlag
- [BMSMS03] BUCHER, Benno; MEIER-SOLFRIAN, Walter; MEYER, Urban; SCHLICK, Sandra: *Statistik*. 2003. Compendio Bildungsmedien AG
- [Erd13] ERDOGMUS, Derman: *Usability-Untersuchung von Liquid Democracy-Systemen*. 2013. Bachelorthesis Heinrich-Heine-Universität Düsseldorf
- [KBBM16] KRAUTHOFF, Tobias; BAURMANN, Michael; BETZ, Gregor; MAUVE, Martin: *Dialog-Based Online Argumentation*. 2016
- [Kol12] KOLLECK, Prof. Dr. B.: *Einführung in die Sozialforschung und Statistik für BA-Studiengänge Soziale Arbeit, Gesundheit und Bildung, Version 2*. 2012. Alice Salomon Hochschule Berlin
- [Kä10] KÄHLER, Wolf-Michael: *Statistische Datenanalyse*. 2010. Springer Verlag
- [Lei04] LEINER, Bernd: *Einführung in die Statistik*. 2004. Walter de Gruyter GmbH und Co KG
- [Rei06] REICHEL, Melanie: *Konzeption, Entwicklung und Usability Evaluation einer Webanwendung für die Verwaltung von Webhosting Leistungen*. 2006. Hochschule für Technik und Wirtschaft Dresden (FH)

Ehrenwörtliche Erklärung

Hiermit versichere ich, die vorliegende Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Alle Stellen, die aus den Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Düsseldorf, 9. September 2016

Daniel Braune

Bitte hier

die Hülle mitsamt CD einkleben

Diese CD enthält:

- Eine *pdf* Version der Bachelorarbeit
- Alle $\text{\LaTeX}$ und Grafik Dateien mitsamt dazugehörigen Skripten, die verwendet wurden
- Alle während der Evaluation angefallenen Auswertungsdaten
- Alle referenzierten wissenschaftlichen Arbeiten
- Alle Testbögen der zwei Tests, des Pre-Tests sowie des Vergleichsbogens