

# Informationsvisualisierung in der Medizin

DIPLOMARBEIT

zur Erlangung des akademischen Grades

**Magister der Sozial- und Wirtschaftswissenschaften**

im Rahmen des Studiums

**Magisterstudium Informatikmanagement**

eingereicht von

**Harald Endl BSc**

Matrikelnummer 0126359

an der  
Fakultät für Informatik der Technischen Universität Wien

Betreuung  
Betreuerin: Ao.Univ.Prof. Mag.rer.soc.oec. Dr.phil. Margit Pohl

Wien, 22.10.2015

\_\_\_\_\_  
(Unterschrift Verfasser/in)

\_\_\_\_\_  
(Unterschrift Betreuer/in)







## Erklärung zur Verfassung der Arbeit

Harald Endl BSc  
Fred-Liewehr-Gasse 21/5  
1130 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

---

Datum, Ort

---

Unterschrift



## Kurzfassung

Gegenstand dieser Masterarbeit war die Evaluierung des Programms „TimeRider“, welches von der Technischen Universität Wien mit Kooperationspartnern entwickelt wurde (Homepage des Projekts „VisuExplore“, 2012).

Das Ziel des Programms ist es, die Entwicklung von medizinischen Parametern über einen Beobachtungszeitraum darzustellen. Die dazu verwendeten Parameter wurden auf Basis fiktiver „Diabetes mellitus“ Patientendaten entwickelt (Homepage des Programms „TimeRider“, 2011).

Die zentrale Fragestellung dieser Arbeit war, ob nach einer kurzen Einarbeitungszeit mittels verschiedener Aufgabenstellungen, ÄrztInnen mit Berufserfahrung ein eigenständiges Arbeiten mit dem Programm möglich wäre. Gegenstand der Untersuchung war somit die Benutzerfreundlichkeit (=Usability) des Programms. Die Untersuchungsmethode der Wahl war „Thinking Aloud“, erstmals im Detail beschrieben in den Arbeiten von Ericsson und Simon (1980, 1993) und in der Arbeit von Boren und Ramey (2000) mit Bezug auf Usability.

Für die Evaluierung von „TimeRider“, wurden Kategorien des Fachgebiets der Informationsvisualisierung herangezogen, welche in den Arbeiten von Forsell & Johansson (2010) und Yi, et al. (2007) beschrieben werden.

Das Ziel der Diplomarbeit war, eine umfangreiche Überprüfung der Software, sowohl auf die didaktische Nutzbarkeit in der medizinischen Ausbildung, als auch auf die Vorzüge und Schwächen unter den Gesichtspunkten der Informationsvisualisierung und der Benutzerfreundlichkeit (*Usability*).

Die Auswertung der gesammelten Daten zeigt, dass ein selbständiges Arbeiten nach einer Einführung und vier selbständig bearbeiteten Aufgaben weitgehend möglich ist. Des Weiteren wurden von allen 10 ProbandInnen Designfehler aufgezeigt, welche die Benutzbarkeit des Programms einschränken.



## Abstract

The aim of this master thesis is the evaluation of the “TimeRider” program, which was developed by the Vienna University of Technology with their cooperation partners (Homepage des Projekts „VisuExplore“, 2012).

The goal of the program is, to develop a graphical representation of medical parameters. The parameters were based on notional “diabetes mellitus” medical records (Homepage des Programms „TimeRider“, 2011).

The focus of this work was to show after a limited training period with different tasks that medical doctors with professional experience can work independently with the program. The aim was to test the usability of the program. The research method of choice was “Thinking Aloud”, first described in the works of Ericsson and Simon (1980, 1993) and in Boren and Ramey (2000) with regard to usability. The master thesis is set in the context and methods of information visualization (Forsell & Johansson (2010) and Yi, et al., 2007).

The objective of this work was primarily to test the software as well for didactical usability in medical education as the advantages and disadvantages under the aspect of information visualization and usability.

The evaluation of the collected data showed that the users are able to work independently with the program after a short introduction and four tasks, which they had to work on independently. Each user also reported design bugs, which limits the utility of the program.



## Danksagung

Diese Diplomarbeit entstand in Zusammenarbeit mit dem Institut für Gestaltungs- und Wirkungsforschung der Technischen Universität Wien und deren „Human Computer Interaction“-Gruppe unter der der Leitung von Univ. Prof. PhD Geraldine Fitzpatrick. Mein besonderer Dank gilt Frau Ao. Univ. Prof. Dr.phil Mag.rer.soc.oec. Margit Pohl, für die allzeit freundliche Unterstützung, sowie Ihrer wertvollen Ratschläge und konstruktiven Kritik. Bei Ulrich Fels BSc und Nikolaus Suchy möchte ich mich für die Mithilfe beim Vorbereiten der Tests bedanken.

Ein ganz besonderer Dank geht an meine beiden Großeltern Friedrich und Johanna Steinhäuser, die immer für mich da waren um meinen Wissensdurst zu stillen und mich motivierten meinen Weg an der Technischen Universität Wien sowie der Medizinischen Universität Wien, als auch an der Universität Wien zu beginnen und auch dementsprechend fortzusetzen. Ebenso möchte ich mich in ganz besonderer Weise bei meiner Familie, meinen Eltern Josef und Regina Endl und meinem Bruder Philipp Endl für die umfassende Unterstützung meines beruflichen Werdegangs bedanken. Ein ganz besonderes Dankeschön möchte ich an meine Lebenspartnerin Frau Jennyfer Gottlieb richten, für Ihre Geduld und Güte sowie die seelische und geistige Unterstützung in meiner Ausbildung und meinem Berufsleben.

Wien, 18.10.2015

Harald Endl BSc



# Inhaltsverzeichnis

<b>Erklärung zur Verfassung der Arbeit</b> .....	<b>i</b>
<b>Kurzfassung</b> .....	<b>iii</b>
<b>Abstract</b> .....	<b>v</b>
<b>Danksagung</b> .....	<b>vii</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Aufbau der Arbeit .....	3
1.2 Methodisches Vorgehen .....	4
1.2.1 Versuchsaufbau .....	4
1.2.2 Methode der Wahl .....	4
1.3 Stand der Forschung .....	5
1.4 Zielsetzung und Motivation .....	7
<b>2 Related Work</b> .....	<b>9</b>
2.1 Kohortenstudien .....	9
2.2 VisuExplore .....	11
2.3 Time Lines (Zeitstrahl) .....	13
2.4 LifeLines .....	16
<b>3 Informationsvisualisierung in der Humanmedizin</b> .....	<b>19</b>
3.1 Allgemein.....	19
3.2 Visualisierung .....	20
3.3 Informationsvisualisierung .....	21
3.4 Wissenschaftliche Visualisierung .....	22
3.5 Medizinische Visualisierung.....	23
3.6 Visualisierung von zeitorientierten Daten .....	25
3.7 Kurze Abgrenzung von Informationsvisualisierung zu Visual Analytics .....	27
<b>4 Animation</b> .....	<b>29</b>
4.1 Allgemein.....	29
4.1.1 Animationsgeschwindigkeit.....	31
4.1.2 Interaktivität.....	32
4.1.3 Zusätzliche Optionen .....	33

4.1.4	Animation vs. statische Darstellung .....	34
4.1.5	Größe der Datensätze.....	34
4.2	Animation in der Trendanalyse .....	35
4.3	Animation in TimeRider.....	36
4.4	Animated Scatter Plot.....	38
<b>5</b>	<b>Usability .....</b>	<b>41</b>
5.1	Allgemein .....	41
5.2	Definition.....	42
5.2.1	ISO 9241.110.....	44
5.3	Jakob Nielsen – Der Usability Papst .....	45
5.3.1	Definition von Usability .....	46
5.3.2	Die zehn Usability Heuristiken.....	47
5.4	Die acht goldenen Regeln des Interface Design.....	49
5.5	Das Informationsvisualisierungs-Mantra .....	52
5.6	Die sieben Interaktionskategorien .....	56
5.6.1	Weitere Interaktionstechniken.....	64
<b>6</b>	<b>Thinking Aloud .....</b>	<b>65</b>
6.1	Thinking Aloud Protokoll nach Boren & Ramey (2000) .....	68
<b>7</b>	<b>Didaktische Verwendungsmöglichkeit .....</b>	<b>81</b>
<b>8</b>	<b>Evaluierung von TimeRider .....</b>	<b>83</b>
8.1	Beschreibung der Untersuchung.....	83
8.2	Beschreibung des Programms TimeRider .....	89
8.3	ProbandInnen.....	96
8.4	Fragestellung .....	98
8.5	Resultate .....	100
8.5.1	Übersicht.....	100
8.5.2	Hypothesen .....	101
8.5.3	Thinking Aloud (Usability aus Sicht der ProbandInnen Aktivitäten).....	107
8.5.4	Usability/Utility Kategorien nach Forsell & Johansson (2010) .....	108
8.5.5	Interaktionskategorien nach Yi, et al. (2007) .....	114
8.5.6	Dokumentierte Hinweise der ProbandInnen.....	118
8.5.7	Abschließende Beurteilung.....	122
<b>9</b>	<b>Zusammenfassung.....</b>	<b>125</b>
9.1	Grenzen der Untersuchung .....	125
9.2	Ausblick.....	126
	<b>Bilderverzeichnis.....</b>	<b>129</b>
	<b>Tabellenverzeichnis .....</b>	<b>131</b>
	<b>Literaturverzeichnis .....</b>	<b>133</b>

---

<b>A</b>	<b>Anhang</b> .....	<b>145</b>
A.1	TimeRider Einschulung .....	147
A.2	Parameterdefinition.....	149
A.3	Liste der Variablen.....	151
A.4	Die vier Aufgaben.....	153
<b>B</b>	<b>Anhang</b> .....	<b>155</b>
<b>C</b>	<b>Anhang</b> .....	<b>157</b>



# 1 Einleitung

In dieser Masterarbeit wird das Programm „TimeRider“ (siehe Abbildung 1) untersucht, welches in Kooperation mit der Technischen Universität Wien, der Donau-Universität Krems und mehreren niederösterreichischen Landesstellen für Gesundheit entwickelt wurde (Homepage des Projekts „VisuExplore“, 2012).

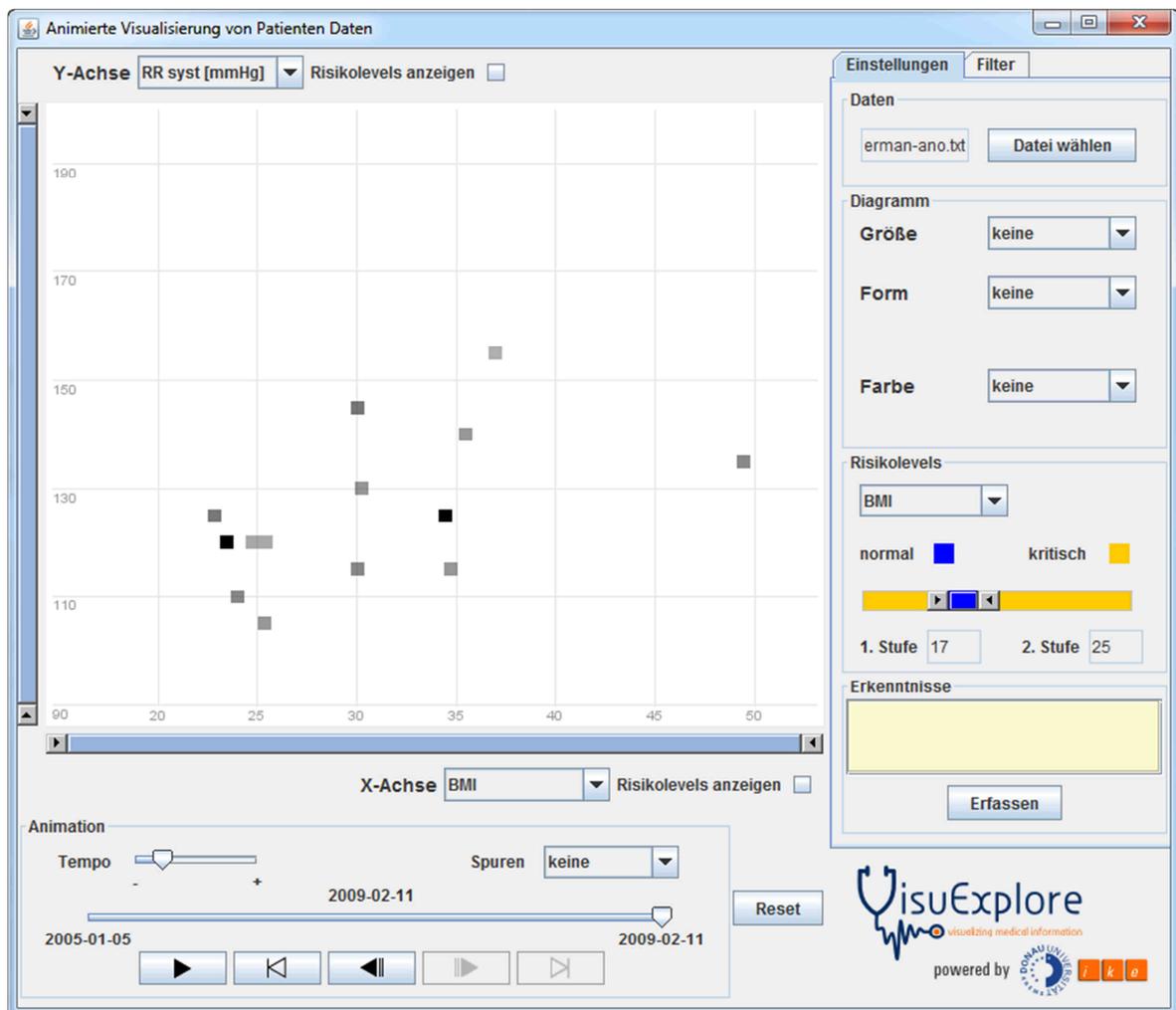


Abbildung 1: User Interface des zu untersuchenden Programms TimeRider

Das Programm ist Bestandteil des „CVASt“ (Centre for Visual Analytics Science and Technology), welches vom Institut für Gestaltungs- und Wirkungsforschung und dem Institut für Software Technologie und interaktive Systeme an der Technischen Universität

Wien ins Leben gerufen wurde. Das Ziel des Zentrums ist es, innovative Methoden zu entwerfen und zu entwickeln, um Informationen mittels interaktiven Visualisierungen und Analysen übersichtlich darzustellen (Homepage des CVASt, 2013).

Die Studie an dem Programm wurde vom „CVASt“ beauftragt und unter der Aufsicht von Frau A.o.Univ.-Prof.in Dr.in Mag.a Margit Pohl und der damaligen Assistentin DI Sylvia Wiltner durchgeführt. Die wissenschaftlichen Ergebnisse wurden in der wissenschaftlichen Arbeit von Rind, et al. (2011) verwendet. Die Untersuchungen und Analysen wurden von den beiden Kollegen Ulrich Fels BSc und Nikolaus Suchy, sowie vom Autor dieser Magisterarbeit selbst durchgeführt.

Im Zuge dieser Arbeit wurde eine „Thinking Aloud“- und eine Interview-Studie an 10 ProbandInnen durchgeführt. Das Ziel der Diplomarbeit war eine umfangreiche Überprüfung der Software bezüglich der didaktischen Nutzbarkeit in der medizinischen Ausbildung, der Vorzüge und Schwächen unter den Gesichtspunkten der Informationsvisualisierung, der Benutzerfreundlichkeit (*Usability*) und des Nutzen eines Programmes (*Utility*).

Die Methode des „Thinking Aloud“ setzt eine selbständige Abarbeitung der gestellten Aufgaben der UserInnen voraus. Ziel ist es, Einsicht in die Gedankengänge und in die Erkenntnisgewinnung der ProbandInnen selbst zu erlangen. Die UserInnen mussten ihre Gedanken offenbaren, indem die getätigten Arbeitsschritte, durch lautes Aussprechen kommentiert wurden. Diese Methode wird im Detail im Kapitel 6 behandelt.

Aufgrund der umfangreichen Arbeiten im Zuge der Untersuchungen, konnten die Ergebnisse nicht nur in einer Diplomarbeit behandelt werden. Es wurde eine zweite Magisterarbeit von Herrn Ulrich Fels BSc mit dem Titel „Usability-Analyse des Programms Animated Scatter Plot“ (Fels, 2015) erstellt. Der Schwerpunkt dieser zweiten Arbeit lag auf Interview-Befragung und behandelt im Detail ergänzende Methoden und Untersuchungsziele des Programms TimeRider.

Das zu analysierende Programm ist somit für beide Arbeiten gleich, aber die Untersuchungsmethoden sind unterschiedlich. Dies wurde vom CVASt vorausgesetzt, um eine wissenschaftliche Analyse der Hypothesen zu gewährleisten. Dadurch kommt es zu wörtlichen und fachlichen Überschneidungen in den beiden Magisterarbeiten (Fels, 2015).

In der Bachelorarbeit des dritten Kollegen Herrn Nikolaus Suchy, werden Verbesserungsvorschläge zu den entdeckten Usability und Utility Fehlern bezüglich des Programms TimeRider diskutiert (Suchy, 2010).

Als Formatvorlage für das Zitieren dieser Arbeit wurde das „Harvard System“ in der Version von Juli 2008 der Anglia Ruskin Universität verwendet (Homepage der Anglia Ruskin Universität, 2008).

## 1.1 Aufbau der Arbeit

Die Gliederung dieser Arbeit beginnt mit einer Einleitung, welche das Projekt und den Aufbau der Arbeit allgemein beschreibt, danach folgen die Kapitel mit den zugrundeliegenden Wissensgebieten. Die Auswertung und Interpretation der gemessenen Daten befinden sich in den beiden letzten Kapiteln, den Abschluss dieser Magisterarbeit bildet der Anhang, welcher aus den verwendeten Unterlagen, Auswertungen und ergänzenden Dokumenten besteht, sowie aus zwei Exkursen.

Die Einleitung beginnt mit einer Beschreibung des Zentrums „CVASt“ und enthält dieses Unterkapitel 1.1 „Aufbau der Arbeit“. Die beiden darauf folgenden Unterkapitel 1.2 und 1.3 beschreiben zuerst das methodische Vorgehen der Untersuchung und dann den aktuellen Stand der Forschung. Den Abschluss der Einleitung bildet der Abschnitt 1.4 in welchem die Zielsetzung und Motivation der Arbeit erläutert werden.

Die nächsten sechs Kapitel enthalten die theoretischen Grundlagen und die in den weiteren Kapiteln verwendeten Begriffe, welche definiert und erläutert werden. Der Anfang ist Kapitel 2 „*Related Work*“, indem die Arbeiten mit ähnlicher Themenstellung kurz beschrieben werden. Darauf folgt im Kapitel 3 eine Vorstellung der Informationsvisualisierung in der Humanmedizin. Das darauffolgende Kapitel 4 enthält eine Einführung in die Animation von Daten und der daraus resultierende Nutzen für visuelle Interpretation von Informationen. In einem Unterpunkt dieses Kapitels wird die Visualisierungsmethode „Animated Scatter Plots“ beschrieben, welche im zu untersuchenden Programm TimeRider verwendet wurde. In Kapitel 5 wird „Usability“ oder auch „Benutzerfreundlichkeit“ erläutert und die dafür aktuell gültigen Regeln für Computerprogramme vorgestellt. Das nächste Kapitel 6 gibt eine detaillierte Beschreibung der Methode der Wahl „Thinking Aloud“ und das Kapitel 7 informiert über die didaktischen Verwendungsmöglichkeiten.

Das vorletzte Kapitel 8 ist in mehrere Unterkapitel gegliedert und enthält die gemessenen Daten, deren Auswertung und Interpretation. Nach einer kurzen Beschreibung der Untersuchung im Kapitel 8.1 folgt im nächsten Unterkapitel 8.2, die Vorstellung des Programms TimeRider und die verwendeten technischen Komponenten. Das Kapitel 8.3 enthält dann eine Beschreibung des ProbandInnenkollektivs und eine Begründung der gewählten Anzahl an Testpersonen. Im Kapitel 8.4 wird im Detail die Fragestellung der Magisterarbeit und die Hypothesen beschrieben. Das Kapitel 8.5 beinhaltet die gemessenen Resultate mit Hinblick auf die Hypothesen, sowie eine Auflistung aller dokumentierten Fehler und Interaktionen. Ebenfalls wird im selben Kapitel die Schwierigkeit der Kategorisierung der dokumentierten Fehler beschrieben.

Das letzte Kapitel 9 enthält die Zusammenfassung und die Grenzen dieser Untersuchung, sowie einen Ausblick auf mögliche zukünftige Arbeiten. Den Abschluss bildet dann, das Bild,- Tabellen- und Literaturverzeichnis, sowie der Anhang der unter anderem die verwendeten Unterlagen, Auswertungen und zwei Exkurse enthält.

## 1.2 Methodisches Vorgehen

### 1.2.1 Versuchsaufbau

Die Hypothesen dieser Magisterarbeit wurden durch vier Aufgabenstellungen überprüft, welche von 10 ProbandInnen gelöst werden sollten. Es konnten 10 fertig ausgebildete ÄrztInnen mit Berufserfahrung und keine HumanmedizinstudentInnen aus verschiedenen humanmedizinischen Fachrichtungen als Testpersonen gewonnen werden. Dies war eine Hauptanforderung des CVASt an diese Arbeit, eine reine ProbandInnen Gruppe aus berufstätigen ÄrztInnen (auch „*Professionals*“) zu bilden. Für bisherige Untersuchungen an ähnlich zu untersuchenden Programmen wie TimeRider, konnten nur HumanmedizinstudentInnen gewonnen werden, da berufstätige ÄrztInnen meist wenig Zeit zur Verfügung haben um an Studien außerhalb Ihres Fachgebiets teilzunehmen (Pohl, et al., 2011). Dem Autor dieser Magisterarbeit gelang es, durch sein privates Netzwerk, Kontakte zu ÄrztInnen die dem Anforderungsprofil entsprachen herzustellen und die gewünschte Anzahl zu motivieren.

Die vier Aufgaben (auch „*Tasks*“) waren für das zu untersuchende Programm TimeRider spezifiziert (siehe Originaldokument im Anhang A.4). Die Tests wurden auf einem Laptop mit einer zusätzlichen externen Computermaus, einem zusätzlichen externen Mikrofon und einem angeschlossenen Flachbildschirm durchgeführt. Zur audiovisuellen Aufzeichnung der Benutzeraktivität, wurde am Laptop das Programm "Windows Media Encoder 9" installiert und verwendet. Die detaillierten technischen Daten und Konfigurationen sind im Kapitel 8.1 beschrieben.

Durch die Mobilität des Laptops, konnten die ProbandInnen an jedem beliebigen Ort befragt werden. Dadurch konnten wir die Bereitschaft zu einer Teilnahme von berufstätigen ÄrztInnen steigern.

### 1.2.2 Methode der Wahl

Die wissenschaftlichen Untersuchungsmethoden, welche zur Erhebung der Rohdaten in dieser Magisterarbeit verwendet worden sind, sind einerseits „Thinking Aloud“ und andererseits Interviews.

Die Methode des „Thinking Aloud“ setzt eine selbständige Abarbeitung der gestellten Aufgaben durch die UserInnen voraus (siehe Kapitel 6). Um Einsicht in die Gedankengänge der ProbandInnen selbst zu erlangen, müssen die UserInnen die getätigten Arbeitsschritte laut aussprechen und kommentieren. Diese Technik wird in der Psychologie und den Sozialwissenschaften verwendet und wurde von Ericsson und Simon (1980) vorgestellt. Clayton Lewis führte dann die Methode in das Forschungsgebiet Informatik unter „Kognitives Benutzeroberflächen Design“ ein (Lewis, 1982). Weitere Details zu dieser Methode werden im Kapitel 6 behandelt.

Die zweite Analysemethode des Programms wurde mittels Interview-Befragung durchgeführt, siehe hierfür die Magisterarbeit von Ulrich Fels BSc mit dem Titel „Usability-Analyse des Programms Animated Scatter Plot“ (Fels, 2015).

Wie bereits erwähnt, wurde zur audiovisuellen Aufzeichnung der Aktivität der BenutzerInnen, das Programm „Windows Media Encoder 9“ verwendet. Dieses erzeugt jedoch keine sogenannte Protokolldatei (auch *Logfile*), eine Datei welche automatisch alle Aktionen erfasst die in einem Computersystem aufgetreten sind (Homepage des freien online Lexikons „Webopedia“, 2013), sondern nur Videodateien, welche Bild und Ton enthielten. Die Inspiration für die Verwendung eines *Logfiles* kam aus Pohl, et al. (2012). Darin ist eine Analyse von Protokolldateien (=Logdateien) beschrieben, die Aktivitäts- und Interaktionsmuster identifiziert, welche die BenutzerInnen während Ihrer Arbeit mit dem System aufwiesen. Aus diesem Grund wurden zwei Logfiles erstellt, eine Datei enthielt alle Aktivitäten und die andere nur die Usability relevanten Ereignisse.

Gotz und Zhou (2008) erwähnen am Rande wie aufwendig die Analyse von Logdateien ist und dass es viel Zeit beansprucht. Es war sehr überraschend für den Autor dieser Magisterarbeit, dass es neben einer Teilzeitbeschäftigung, ein Jahr dauern sollte bis alle Arbeiten an der Logdatei abgeschlossen wurden. Der Aufwand wurde unterschätzt, da neben der mehrmaligen Durchsicht der Rohdaten und der Videodateien, um Fehlerquellen auszuschließen, zusätzlich noch die Erstellung der Logfiles, die Aufbereitung der Daten und die anschließende Analyse vollständig manuell durchgeführt werden mussten. Wie zuvor erwähnt war im Aufzeichnungsprogramm kein Automatismus für die Erstellung einer Protokolldatei verfügbar.

### 1.3 Stand der Forschung

Untersuchungen und Studien über Informations- oder Graphen Visualisierungsmethoden beschäftigen sich laut Friel, et al. (2001) meist mit kontrollierten Datensätzen und simplen Graphen. Die Erforschung von multivariaten (= mehrere Variablen betreffend, siehe

Homepage des „Duden“, 2013a), realen Datensätzen, welche komplexe Darstellungsme-  
thoden erfordern, wie in dieser Magisterarbeit durchgeführt, sind laut Trafton, et al. (2000)  
die Ausnahme.

Als Grundlage dieser Magisterarbeit kamen zwei verwandte Forschungsgebiete in Frage:  
die Informationsvisualisierung und Visual Analytics, wobei beide jeweils unterschiedliche  
Visualisierungsansätze beschreiben. Als Grundlage für diese Magisterarbeit diente jedoch  
die Informationsvisualisierung (siehe Kapitel 3.7).

Die Informationsvisualisierung ist der Transformationsprozess von Daten, Informationen  
und Wissen in eine visuelle Form unter Nutzung der natürlichen visuellen Fähigkeiten des  
Menschen (Gershon, et al., 1998). Entstanden aus verschiedensten Fachgebieten, wie  
Mensch-Computer-Interaktion, Informatik, Grafik, visuelles Design (Gestaltung), Psycho-  
logie und Geschäftspraktiken (-methoden); wurde die Informationsvisualisierung unerläss-  
lich in diversen Einsatzgebieten, wie wissenschaftliche Forschung, digitale Bibliotheken,  
„Data-Mining“ (Extrahieren von Informationen aus großen Datenbeständen, siehe Home-  
page des deutschen Wörterbuchverlags „Pons“, 2013), Finanzdatenanalyse, Marktstudien,  
Herstellungsproduktionskontrolle und Pharmaforschung (Bederson & Shneiderman, 2004).  
In Kapitel 3 dieser Arbeit wird näher auf das Gebiet der Informationsvisualisierung einge-  
gangen.

Das zweite Forschungsgebiet ist Visual Analytics, welches laut Keim, et al. (2004) eben-  
falls die Vorteile aus unterschiedlichen Forschungsgebieten interdisziplinär verbindet. Bei  
dieser Methode wird versucht, aus großen und komplexen Datensätzen Erkenntnisse zu  
gewinnen, in dem die Fähigkeiten des Menschen, wie schnelles Erkennen von Mustern  
oder visuelles Erfassen von Trends. Diese werden mit automatischer Datenanalyse kombi-  
niert und durch Interaktionsmechanismen können Daten visuell exploriert und Erkenntnis-  
se gewonnen werden. Der Begriff „Visual Analytics“ wurde 2004 eingeführt und ein Jahr  
später in dem Buch "Illuminating the Path" von Thomas & Cook (2005 zitiert in Keim, et  
al., 2004) beschrieben (Keim, et al., 2004). Auf der Homepage des CVAIST (2013) wird  
visuelle Analytik (*visual analytics*) als ein Forschungsgebiet beschrieben, welches als Ziel  
hat, neue Methoden und Technologien für visuelle Analysen von Datenmengen aus ver-  
schiedenen Applikationen, bei Bedarf auch für kumulierte Datenströme und somit größe-  
ren Datenmengen, zu entwickeln. Dieses Forschungsgebiet dient jedoch nicht als Grundla-  
ge für diese Arbeit (siehe Kapitel 3.7).

Digitale Krankenakten stellen an Systementwickler spezielle Herausforderungen, wie Inf-  
rastruktur- und Datenschutzthemen welche berücksichtigt werden müssen, bevor ÄrztIn-  
nen die Akten verwenden dürfen. Für die Arbeit (z.B. Anamnese) von ÄrztInnen sollte  
mobile und nicht intrusive Hardware zur Verfügung gestellt werden, anstatt immobile, an  
einen Schreibtisch gebundene und sperrige Workstations. Ein Einsatz von langen zu durch-

suchenden Listen, endlosen zu navigierende Menüs und lange Dialoge können dazu führen, dass die BenutzerInnen das Design ablehnen. Daher ist der nächste logische Schritt einen angemessenen Aufwand für ein angepasstes Benutzeroberflächendesign (*User Interface Design*) zu investieren, nur dann können die Herausforderungen erfolgreich gelöst werden (Dewey, 1993 zitiert in Plaisant, et al., 1998). Neue Techniken welche zusammenfassen, filtern und eine große Anzahl von Informationen verständlich machen können, sind der Beweis dafür, dass ein schneller Zugriff auf die benötigten Daten durch ein sorgfältiges Design möglich ist (Plaisant, et al., 1998).

## 1.4 Zielsetzung und Motivation

Die Zielsetzung und Motivation dieser Magisterarbeit war es, im Einklang des didaktischen Anspruches an diese Arbeit, den Lerneffekt bei ÄrztInnen zu untersuchen, welche mit dem Programm TimeRider nach einer kurzen Einführung und Einarbeitungsphase selbstständig arbeiteten.

TimeRider kann von ÄrztInnen in der Qualitätskontrolle und in der klinischen Forschung eingesetzt werden, in Verbindung mit deren klinischen Arbeit. Ein Beispiel dafür wäre die Überprüfung des Erfolgs einer Therapie oder ob diese effektiver ist als vergleichbare Therapien. ÄrztInnen sind daran interessiert ob der Gesundheitszustand sich verschlechtert oder verbessert. Von Interesse ist auch eine Beziehung herzustellen, zwischen verschiedenen Variablen und deren Veränderungen über die Zeit (z.B. Blutzucker und Gewicht). Verschiedene Vergleichs- und Filteroptionen in TimeRider sind für die klinische Forschung interessant, da Patientenkollektive meist große Gruppen von Personen (=Kohorten) sind und über die Zeit durch die graphische Darstellung visuell einfach verglichen werden können (Rind, et al., 2011 & Weiß, 2013, p.252).

Darüber hinaus wurden Fehler im Design des Programms hinsichtlich „Usability“ (= Benutzerfreundlichkeit eines Programmes) untersucht und beschrieben. Diese Erkenntnisse sollen in zukünftige Aktualisierungen (*Updates*) des Programmes einfließen.



## 2 Related Work

Dieses Kapitel beschreibt kurz Arbeiten mit ähnlicher Themenstellung. In vier Unterkapiteln werden zum einen Kohortenstudien, das Projekt „*VisuExplore*“, die Darstellungsform „*TimeLines*“ und die Visualisierungsumgebung „*LifeLines*“ vorgestellt.

### 2.1 Kohortenstudien

Im Gesundheitswesen sind die Analyse und die Erforschung von PatientInnen durch Kohortenstudien, vor allem für Gesundheitsdienstleister oder die klinische Forschung von Interesse. Diese Studien zeigen nicht nur die Entwicklung von einen oder mehreren Variablen über die Zeit, sondern auch die Beziehung der Variablen untereinander und den daraus entstehenden Wechselwirkungen innerhalb eines gewählten Zeitraums (Rind, et al., 2011).

Im Detail beschreibt eine Kohortenstudie mindestens zwei oder auch mehrere Patientenuntersuchungsgruppen beginnend mit der Exposition (Patienten sind Risikofaktoren ausgesetzt) bis zum Resultat (siehe Abbildung 2). Diese Art von Studien kann durch eine prospektive (vorausblickende) Analyse, beginnend von der Gegenwart in die Zukunft, oder eine retrospektive (rückblickende) Analyse, beginnend von der Vergangenheit in die Gegenwart, durchgeführt werden (Grimes & Schulz, 2002).

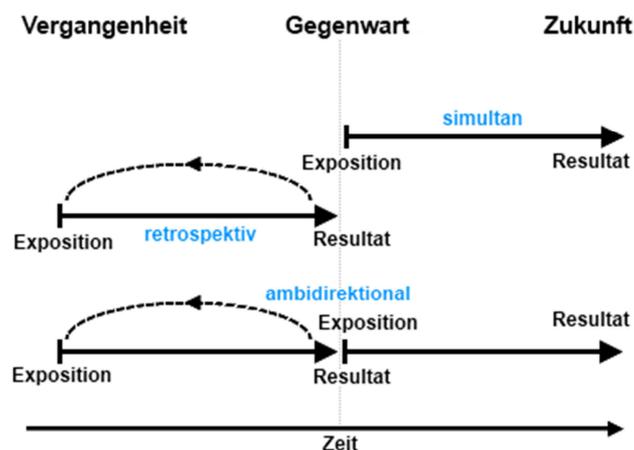


Abbildung 2: Zeitlicher Ablauf von Kohortenstudien (angelehnt an Grimes & Schulz, 2002).

In ihrer einfachsten Form werden bei diesen Studien die Daten einer Patientengruppe, welche exponiert war, mit der einer Untersuchungsgruppe, welche diesem nicht ausgesetzt waren (=Kontrollgruppe), verglichen. Bei einem höheren oder niedrigeren Unterschied der Resultate der Risikogruppe, als der Resultate der Kontrollgruppe, gilt der Zusammenhang zwischen Risikofaktor (Exposition) und Gesundheitszustand (Resultat) als offensichtlich (Grimes & Schulz, 2002).

Die Herausforderung bei Patientendaten von Kohortenstudien sind meist unregelmäßig gesammelte Daten, Datenverschleiß und der Vergleich von Datensätzen welche verschiedene Zeitabschnitte abdecken. Die Daten verschiedener Patienten sind üblicherweise unabhängig voneinander gesammelt worden, im Kontrast zu Rosling (Rosling, 2007 zitiert in Rind, et al., 2011) wo Daten jährlich erhoben wurden. Eine weitere Herausforderung ist die Veränderung, meist Senkung, der Gültigkeit von medizinischen Parametern über längere Zeiträume. Dadurch kann sich ein zusätzliches Risiko ergeben, wenn Daten von zwei verschiedenen voneinander unabhängigen Arbeiten zur weiteren Berechnung oder Verarbeitung verwendet werden würden, da die Verminderung der Gültigkeit der medizinischen Parameter verborgen bliebe. Ebenfalls eine Herausforderung ist die visuelle Darstellung von sequentiell behandelten Patienten, diese sollten parallel dargestellt werden, damit ein visueller Vergleich ermöglicht wird (Rind, et al., 2011).

Das in dieser Magisterarbeit untersuchte Programm TimeRider soll den BenutzerInnen nun helfen die Zusammenhänge zwischen Exposition und Resultat zu analysieren. Mittels der Konzepte der explorativen Datenanalyse (statistische Analyse von wenig oder unbekanntem Daten) und visuellen Methoden, sollen sich die ÄrztInnen einen Überblick über die gemessenen Daten und deren Wechselwirkungen verschaffen können. Dies kann zu neuen Hypothesen oder neuen Einblicken führen, welche bei der Verwendung von klassischen statistischen Methoden (Regressions-, Diskriminanzanalyse) möglicherweise nicht eindeutig oder gar nicht entdeckt worden wären. Die Verwendung von Visualisierungsmethoden helfen jede(r) BenutzerIn und nicht nur ExpertInnen dabei computergenerierte Ergebnisse zu interpretieren. Für ÄrztInnen ist es im Detail z.B. von Interesse, einen Überblick zu gewinnen, wie sich der Patientengesundheitszustand über die Zeit verändert oder auch nicht. Dieses Datenverhalten würde dann Rückschlüsse ermöglichen, ob eine bestimmte Patientengruppe sich anderes verhält, als die restlichen untersuchten Patientenkollektive. Im Detail müsste dann ausgewertet werden, ob dieses Datenverhalten aufgrund einer Behandlungsmethode, aufgrund der Lebensweise des Patienten oder aus einer Kombination aus beidem resultiert (Rind, et al., 2011).

Damit ÄrztInnen die Zusammenhänge in TimeRider schnell und einfach erkennen und begreifen können, wurde als Visualisierungsmethode ein animiertes Streudiagramm (=„*Animated Scatter Plots*“) gewählt. Im Kapitel 4.4 wird diese visuelle Methode im Detail

erläutert. Sie wurde vor allem gewählt um die Beziehungen zwischen den verschiedenen Variablen zu erforschen, welche von medizinischen Interesse sind (Rind, et al., 2011).

## 2.2 VisuExplore

Das Projekt VisuExplore wurde von Frau Ao.Univ.Prof.in Mag.a Dr.in Silvia Miksch initiiert und dauerte von März 2008 bis März 2011 und wurde von den folgenden Projektpartnern realisiert: Donau Universität Krems/Institut für Informations- und Wissensmodellierung (Homepage des „DIKE“, 2011), Technische Universität Wien/Institut für Gestaltungs- und Wirkungsforschung/Forschungsgruppe Human Computer Interaction (Homepage der „HCI“, 2013), NÖ Gesundheits- und Sozialfonds (Homepage des „NÖGUS“, 2013), NÖ Landeskliniken-Holding (Homepage der „LKNOE“, 2013), Landeskrankenhaus Krems (Homepage des „LKNOE-UK Krems“, 2013), systema Human Information Systems (Homepage der Firma Systema, 2013) (Homepage des Projekts „VisuExplore“, 2012).

Das Ziel von VisuExplore war es eine flexible und interaktive Visualisierungsumgebung für zeitorientierte medizinische Daten und Informationen zu implementieren. Diese Methoden und Tools sollen in weiterer Folge bei der Datenanalyse, den Auswertungsaufgaben und den medizinischen Behandlung, die ÄrztInnen und das klinische Personal unterstützen, sowie die Qualität der Pflege verbessern (Homepage des Projekts „VisuExplore“, 2012).

Der Einsatz von moderner Technologie im klinischen Alltag, führt zu einem massiven Anstieg der Quantität und der Komplexität von elektronisch erfassten medizinischen Daten und Informationen. Interaktive visuelle Darstellungen können dabei helfen den Informationsüberschuss zu vermeiden und die Informationen von multiplen heterogenen Datenquellen besser verständlich zu machen, im speziellen im medizinischen Bereich der zeitorientierten Daten. Diese Darstellungen umfassen Änderungen des Gesundheitszustandes von Patienten, therapeutische Maßnahmen, Interventionen und Messungen, sowie Trends, Muster und Beziehungen zwischen Variablen über die Zeit (Homepage des Projekts „VisuExplore“, 2012).

Der zentrale Punkt bei Visualisierungen von medizinischen Daten ist nicht die Präsentation oder das graphische „Rendering“, (laut Homepage des „Duden“, 2015a, bedeutet „Rendering“: „Vorausberechnung [...] am Computer mithilfe einer dreidimensionalen virtuellen Darstellung [...]“), sondern die Gewinnung von neuen Einblicken und Wissen durch intensive und aktive Untersuchung der gesammelten Daten und Informationen. Statische Visualisierungen unterstützen diese Analyseprozesse nicht, da sie nur eine passive Ansicht bieten. Visuelle Untersuchungen benötigen flexible Methoden zur Interaktion und Manipulation mit den Daten und der Darstellung (Homepage des Projekts „VisuExplore“, 2012).

Im Projekt VisuExplore entstanden zwei Forschungsprototypen „*VisuExplore-Prototyp*“ und TimeRider für zwei jeweils verschiedene Anwendungsszenarien, wobei der VisuExplore-Prototyp in diesem Kapitel kurz beschrieben wird und TimeRider ausführlich im Kapitel 8.2 behandelt wird. Die Entwicklung der zwei Prototypen im Projekt VisuExplore wurde durch eine Förderung im Rahmen des Brückenschlagprogramms der Österreichischen Forschungsförderungsgesellschaft (FFG) ermöglicht.

Der VisuExplore-Prototyp ist eine flexible Interaktions- und Visualisierungsumgebung für elektronische Krankenakten eines einzelnen Patienten (Homepage des Projekts „*VisuExplore*“, 2012) und war ebenfalls auf interaktive visuelle Untersuchung von medizinischen Daten ausgelegt, hatte jedoch noch nicht das überarbeitete Design der Animated Scatter Plots implementiert. Der Prototyp war bereits für folgende Hauptanforderungen designt: einfache Benutzeroberfläche (*user interface*), flexible Auswahlmöglichkeiten verschiedenster medizinische Parameter, zeitorientierte Daten, mehrere Patienten, Interaktivität (Unger, 2014).

Abbildung 3 zeigt die zeitlich ausgerichtete Visualisierung des Prototyps von verschiedenen klinischen Parametern für einen zufällig gewählten Diabetes mellitus Patienten.

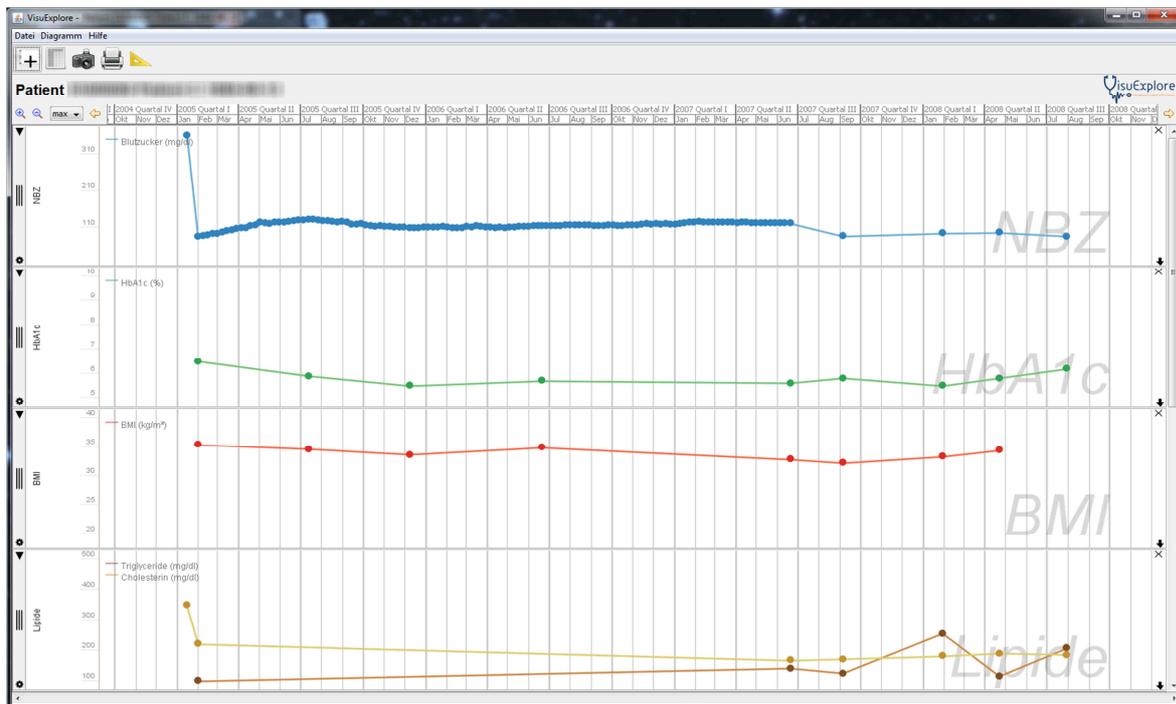


Abbildung 3: User Interface des VisuExplore-Prototypen

Das Ziel war es, bereits in den Prototypen eine Integration verschiedenster Visualisierungs- und Interaktionstechniken vorzunehmen, wobei auch eine einfache Erweiterung von weiteren Techniken möglich sein sollte. Somit können die UserInnen vier verschiedene Visualisierungstechniken für jeden Parameter auswählen, diese sind: Ereignisdiagramm, Liniendiagramm, Balkendiagramm, Streudiagramm und Zeitleisten. Ebenfalls können Parameter

jederzeit hinzugefügt oder entfernt werden. Es werden auch Interaktionstechniken angeboten, wie z.B. die Größenänderung der Parameteransicht oder die Veränderung der Zeitachse durch zoomen und bei hineingezoomten Zeitabschnitten durch klicken und ziehen mit der linken Taste der Computermaus (=Maustaste) (Unger, 2014).

## 2.3 Time Lines (Zeitstrahl)

Laut der Homepage des deutschen Duden, bedeutet „*Time Lines*“ zu Deutsch „*Zeitstrahl*“: „*grafische Darstellung einer geschichtlichen Entwicklung als Linie mit Markierungen für Abschnitte oder Ereignisse*“ (Homepage des „Duden“, 2015b). Tufte (1983 zitiert in Neubauer, 2012) beschreibt Time Lines als eine häufig und mächtige Form von Grafikdesign (Plaisant, et al., 1998). Ein berühmter Zeitstrahl ist die Darstellung von Charles Joseph Minard, welche die Verluste an Soldaten, die Truppenbewegungen und die Temperaturen im Laufe von Napoleons Russlandfeldzugs 1812 abbildet (Marey, 1885, p.72), siehe hierfür Abbildung 4. Die breite Linie welche links beginnt zeigt die Mannstärke der Armee am Beginn des Feldzuges, welcher an der polnisch-russischen Grenze begann und nach rechts immer kleiner wird aufgrund der hohen Verluste. Am unteren Rand sind in Korrelation die Datumsangaben und Temperaturen, denen die Armee ausgesetzt war, eingetragen. Man sieht auf einen Blick, dass die breite Linie immer schmaler wird, welche die Truppenstärke abbildet. Dies lag daran, dass je länger der Feldzug dauerte und umso näher der Winter kam, desto kälter wurde es und die Verluste stiegen, da die Soldaten einfach erfroren (Homepage von „nowscape.com“, 2015).

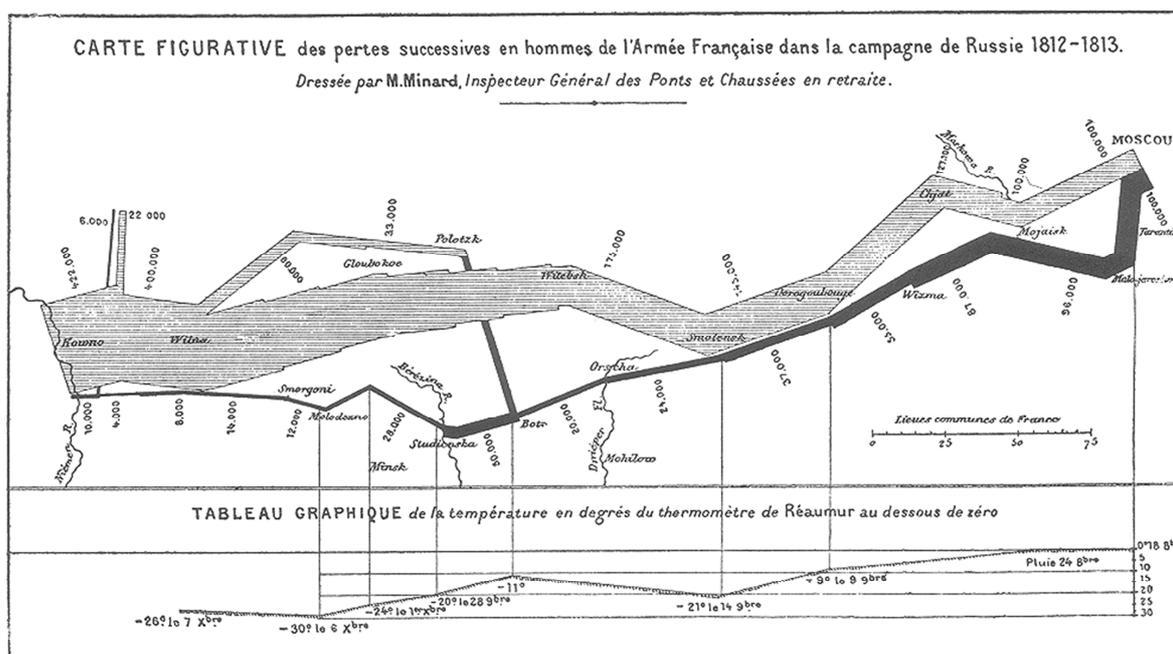


Abbildung 4: Napoleons Russlandfeldzug dargestellt von Charles Joseph Minard (Homepage von „Info-Vis.net“, 2015)

Eine Visualisierung für Krankenakten mit Time Lines wurde in Powsner & Tufte (Powsner & Tufte, 1994 zitiert in Plaisant, et al., 1998) vorgestellt. Sie beschreiben darin, wie sich die Patientenakten von handschriftlichen Notizen, Formularen und Ausdrucken von der Antike, über die 1960ern und bis heute gehalten haben. Der Grund für den Erfolg der klassischen „Hardware“ Aufzeichnungen in der Medizin ist bis heute, dass die meisten vorgestellten „Verbesserungen“ großen Aufwand mit sich bringen würden, im krassen Gegensatz zum einfachen Notizen schreiben und Anweisungen geben (Powsner & Tufte, 1994).

Das präsentierte Visualisierungsdesign ist eine graphische Zusammenfassung eines Patientenstatus, als Alternative zur traditionellen Krankenakte mit Befunden. Numerische Daten sind auf einer Seite abgebildet und zeigen auf einen Blick den Verlauf einer Krankheit und deren Reaktion auf medizinische Behandlungen. Die neuesten Werte sind sowohl graphisch als auch numerisch dargestellt (siehe Abbildung 5). Notizen und medizinische Bilder in Briefmarkengröße sind in eine passende Form für hochauflösende Computerbildschirme oder Laserdruckseiten gebracht worden. Somit entstand eine detaillierte einseitige Zusammenfassung des Patientenstatus (siehe Abbildung 6). Eine Anamnese muss nicht einzig und allein ein nüchterner Bericht des Lebens und der Krankheit des Patienten sein. Eine grafische Zusammenfassung ersetzt jedoch die traditionellen Aufzeichnungen nicht mehr, als eine Landkarte ein Telefonbuch ersetzt. Ähnlich wie eine Landkarte stellt eine Grafik einen Überblick auf die Krankheit eines Patienten zur Verfügung, wobei ÄrztInnen auf einen Blick sehen können, wie diverse Befunde und Behandlungen miteinander in Beziehung stehen (Powsner & Tufte, 1994).

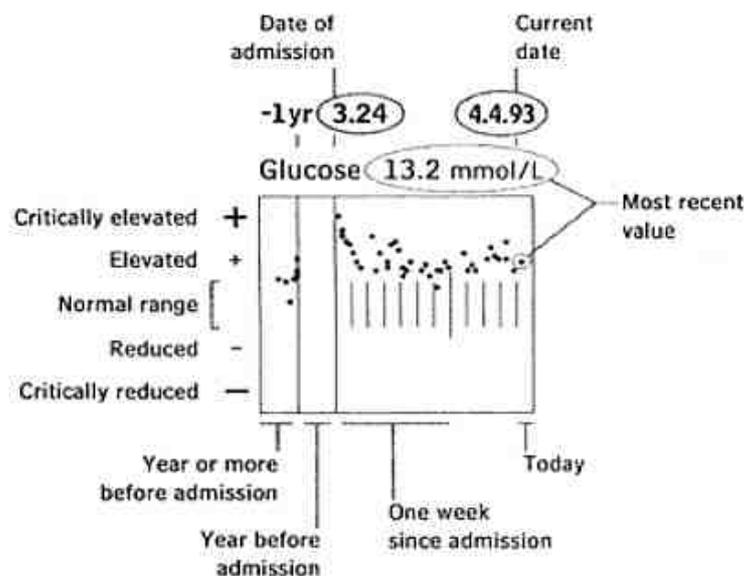


Abbildung 5: Darstellung des Powsner & Tufte „all data chart“ (Homepage der „IFS“, 2006).

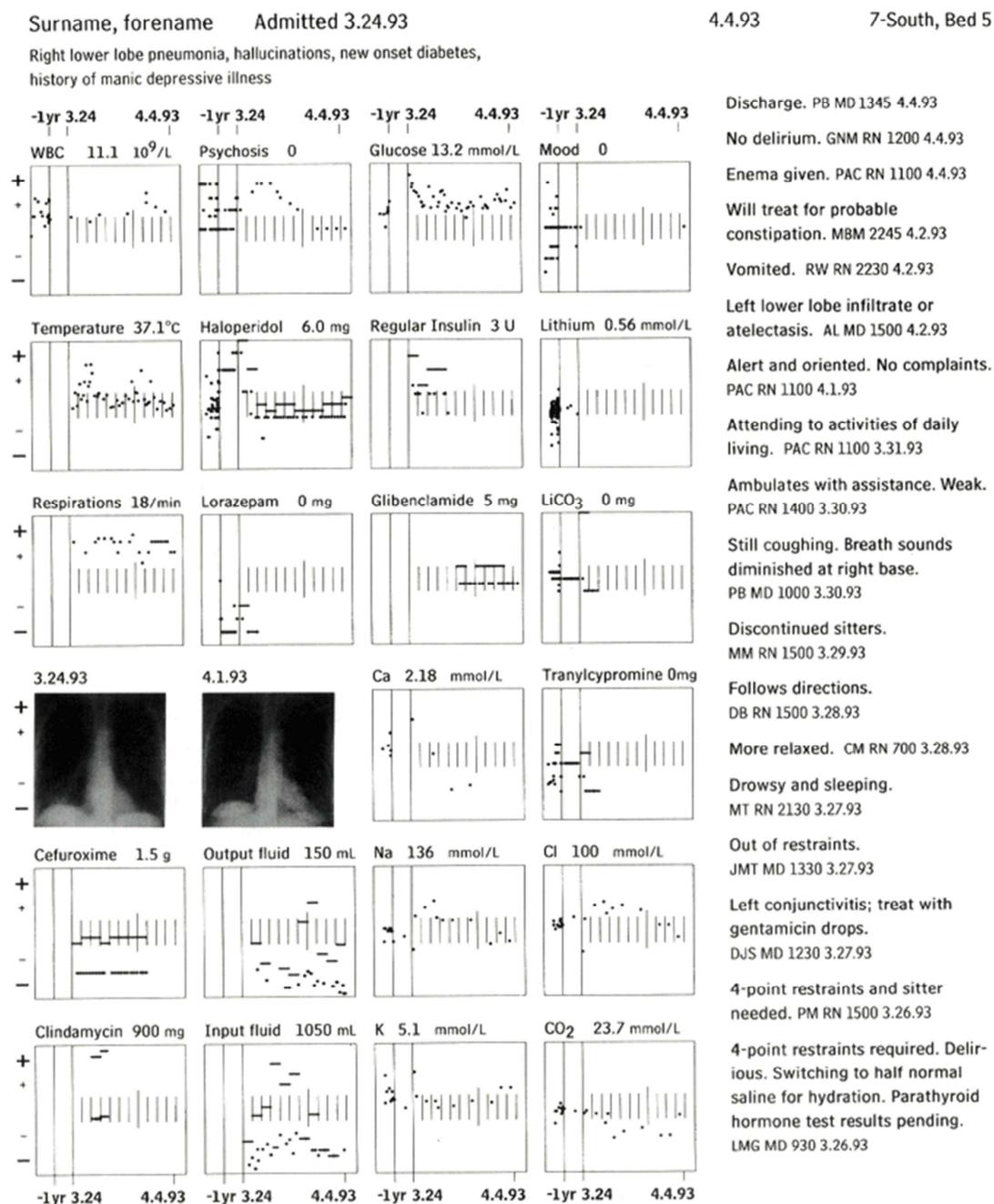


Figure 2: Graphical summary of patient status

Notes on right show initials and professional qualification (MD = medical doctor, RN = registered nurse), time, and date (American format). — = period of dosing.

measurements and treatments are possible. Different time scales and vertical scales would be appropriate for cyclical syndromes (eg, malaria) or slowly changing findings (eg, tumour markers).

Full-page computer displays can show figure 2 as printed here. Although providing the same information as in figure 2, full-page screens are coarser because of larger pixel size compared with the laser-printer's dot. Computer displays

## 2.4 LifeLines

„LifeLines“ bietet eine generelle Visualisierungsumgebung für personenbezogene Historien, wie zum Beispiel klinische Patientenakten. Eine in Java erstellte Benutzeroberfläche (siehe Abbildung 7) wurde in Plaisant, et al. (1998) vorgestellt. LifeLines verwendet die Grafikdesignmethode Time Lines (siehe vorheriges Kapitel 2.3) und ermöglicht einen Überblick einer computerisierten Patientenakte. Bei der Repräsentation können Diagnosen, Testergebnisse, Komplikationen oder die Medikation durch grafische Punkte oder horizontale Linien angezeigt werden. Durch Zoomen können mehr Details untersucht werden, Linienfarbe und –stärke weisen auf Zusammengehörigkeit oder Signifikanz hin. Die visuelle Anzeige dient als Schnittstelle mit direktem Zugriff auf die Daten (Plaisant, et al., 1998).

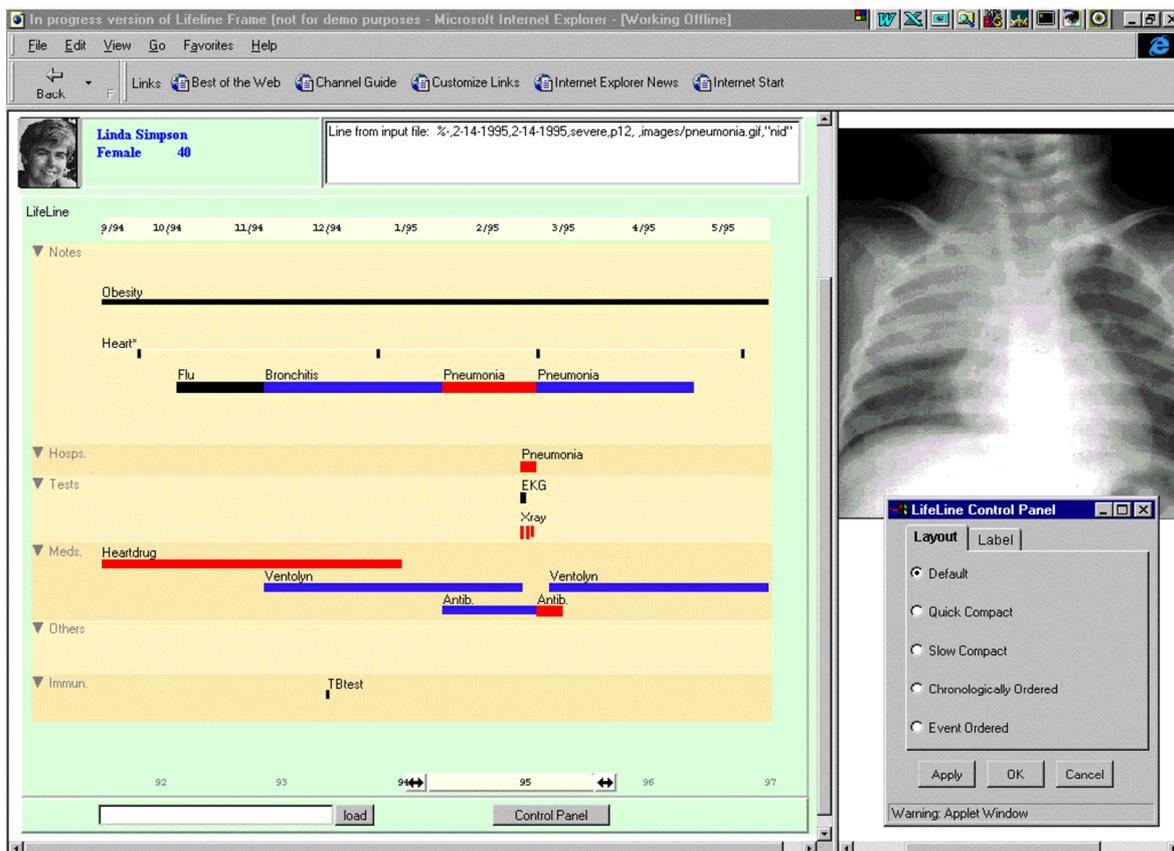


Abbildung 7: User Interface einer elektronischen Patientenakte (Plaisant, et al., 1998, p.3).

Obwohl große Anstrengungen unternommen werden, Standards für die Datensammlung von Patientenakten zu entwickeln, liegt wenig Fokus auf dem Design von passenden Visualisierungs- und Navigationstechniken zur Darstellung und Erforschung von Patientenhistoriendaten. Eine intuitive Möglichkeit um Historien darzustellen ist die Verwendung von grafischen Zeitreihen und wurde in Plaisant & Rose (Plaisant & Rose, 1996 zitiert in Plaisant, et al., 1998) vorgestellt. Die konsistente lineare Zeitskala erleichtert zu vergleichen und Beziehungen zwischen den angezeigten Mengen zu erkennen. Die Daten können

grafisch als Time Lines dargestellt werden, um die Entwicklung der quantitativen Daten über die Zeit zu veranschaulichen. Schnittstellen mit einer hohen Interaktivität verwandeln den Bildschirm in ein strukturiertes Menü mit direktem Zugriff auf jegliche Informationen, welche notwendig sind um eine Patientenakte zu rezensieren (Plaisant, et al., 1998).

LifeLines basiert auf Time Lines und nutzt den Vorteil der menschlichen Fähigkeit, visuell viel Information auf einen Bildschirm analysieren zu können und erleichtert dadurch den Zugriff auf Details in den Patientenakten. Ein solches Interface hat einen signifikanten Einfluss auf die Usability von elektronischen Krankenakten. Ein Beispiel dafür ist die Krankenakte in Papierform, in der das Wichtigste die Befunde und die zusätzlichen Verlaufsnotizen mit den aktuellsten Daten sind. Die Verlaufsnotizen fassen den aktuellen klinischen Status des Patienten zusammen und koordinieren diese Details zeitlich und sachlich. Darüber hinaus enthalten sie Kommentierungen zu den Daten und den Ergebnissen. LifeLines enthält eine einfache und leicht zugängliche Ansicht von zeitlichen und kausalen Beziehungen, dadurch wären die Verlaufsnotizen nur noch eine Kommentarsammlung. Somit müsste bei einer digitalisierten Krankenakte, welche LifeLines anwendet, ein neues Format oder andere Funktionen für die Verlaufsnotizen gefunden werden (Plaisant, et al., 1998).



## **3 Informationsvisualisierung in der Humanmedizin**

Dieses Kapitel beschreibt das Forschungsgebiet Informationsvisualisierung, welches als Grundlage dieser Arbeit dient. Im Unterkapitel 3.1 ist eine allgemeine Definition sowie eine kurze Historie beschrieben. Die darauffolgenden Unterkapitel 3.2, 3.3, 3.4 und 3.5 beschreiben verschiedene Teilbereiche der Visualisierung. Das Unterkapitel 3.6 beschreibt die Visualisierungen von zeitorientierten Daten und den Abschluss bildet das Unterkapitel 3.7 mit einer kurzen Abgrenzung von Informationsvisualisierung zu Visual Analytics.

### **3.1 Allgemein**

Informationsvisualisierung in der Humanmedizin ist ein spezielles Forschungsgebiet innerhalb der Informationsvisualisierung und hat sich durch den verstärkten Einsatz von Computern und Visualisierungen in der Humanmedizin entwickelt.

Die Entstehung der allgemeinen Informationsvisualisierung geht zurück auf einfache Visualisierungen die bildliche Darstellungen waren. Die bis dato ältesten, bekannten und nachgewiesenen Bilder befinden sich in der El-Castillo-Höhle (siehe Abbildung 8) in Spanien und sind ca. 40.000 Jahre alt. Die Bilder entstanden vermutlich, indem die Hand auf die Höhlenwand gelegt wurde und Farbpigmente direkt über die Hand und somit auch auf die Wand gepustet wurden (Valdes & Bischoff, 1989).

Schon damals wurden Sie vermutlich eingesetzt um Sachverhalte zu veranschaulichen. Seit diesen Höhlenmalereien entwickelten sich Darstellungen vor allem in der Astronomie, der Meteorologie und der Kartographie und wurden hauptsächlich zur Unterstützung der Seefahrt und des Militärs im 17. und 18. Jahrhundert eingesetzt. Durch die Erfindung der Computertechnik und Computergraphik, im 20. Jahrhundert konnten sich daraus verschiedenste Visualisierungsmethoden entwickeln (Schumann, 2004).



Abbildung 8: Bild der Höhlenmalerei in der El-Castillo-Höhle (vgl. Pike, et al., 2012, *Supplementary Materials*, p.14)

## 3.2 Visualisierung

Laut der Homepage des “ACM SIGGRAPH Education Committee” (2007) ist die Definition für Visualisierung: *„Visualisierung ist die Bildung von visuellen Bildern, der Vorgang oder Prozess der Interpretation von visuellen Ausdrücken oder eine Transformierung in visuelle Formen. Des Weiteren wird Visualisierung als ein Werkzeug oder eine Methode beschrieben, um in einem Computer eingegebene Bilddaten zu interpretieren und um Bilder aus komplexen mehrdimensionalen Datensätzen zu generieren.“* (Homepage des “ACM SIGGRAPH Education Committee”, 2007).

Der zentrale Punkt der Visualisierung ist das intuitive Erkennen, Verstehen und Bewerten wesentlicher Attribute in Datenmengen, als Unterstützung dienen statistische Berechnungen. Aus automatisierten analytisch-statistischen Berechnungen mit visuellen Darstellungen entstand daraus das Forschungsgebiet „Visual Data Mining“ (Keim, et al., 1997), welches bereits fortlaufend erforscht und erschlossen wird (Schumann, 2004). Wobei „Data

Mining“ grundsätzlich eine statistische und systematische Methode bezeichnet, mit dem Ziel neue Muster in einem Datenbestand zu erkennen (Fayyad, et al., 1996).

In Spence (2001 zitiert in Mazza, 2009, p.7) wird eine breite Verwendung des Begriffs Visualisierung beschrieben und dieser als eine Aktivität dargestellt, an der Menschen durch innere Gedankenkonstruktionen beteiligt sind. Die Visualisierung stellt unter diesen Gesichtspunkt und unterstützt durch externe visuelle Repräsentationen, eine kognitive Aktivität dar, wodurch der Mensch eine interne mentale Repräsentation der Welt erzeugen kann (Spence 2001 und Ware 2004 zitiert in Mazza, 2009, p.7).

Computer unterstützen den Visualisierungsprozess durch Visualisierungs-Werkzeuge, dies ist vor allem durch laufend sinkende Kosten für PCs bei steigender Leistung möglich geworden. Visualisierung ist und bleibt jedoch ein gedanklicher Vorgang des Menschen, trotz der heute möglichen Computerunterstützung (Mazza, 2009).

### 3.3 Informationsvisualisierung

In der Informationsvisualisierung werden Daten visualisiert, welchen keine reale Gegenstände, Strukturen oder einfache Abbildungen zu Grunde liegen, meistens enthalten die Daten keine grafischen oder räumlichen Darstellungen mehr. Es werden dabei computer-gestützte, interaktive, visuelle Repräsentationen von abstrakten, nicht physikalisch-basierten Daten verbildlicht, bzw. visualisiert. Das Ziel ist es diese Daten auf einen Blick verstehen zu können und leichter bearbeitbar zu machen (Card, et al., 1999).

Die Bezeichnung „*Informationsvisualisierung*“ wurde von den Wissenschaftlern am *Xerox Palo Alto Research Center* (Homepage des „*Xerox PARC*“, 2013) Ende der 1980er geprägt. Das Ziel war einen neuen Wissenszweig zu generieren, welcher das Ziel hat: visuelle Artefakte zur Wahrnehmungserweiterung zu erzeugen (Mazza, 2009). Dieser neue Forschungszweig stützt sich auf die Erkenntnis, dass der Mensch ca. 75 Prozent der Information aus seiner Umwelt durch visuelle Signale wahrnimmt. Akustischen Signalen wird ein Anteil von ca. 13 Prozent zugeordnet, den Übrigen insgesamt 12 Prozent (Hölscher & Suchanek, 2011). Dabei wird unter „*Information*“ der (erfolgreiche) Transfer von Wissen verstanden (Homepage der Universität des Saarlandes, 2010).

Laut der Definition nach Card, et al. (1999), ist unter Informationsvisualisierung die Benutzung von computergestützten, interaktiven, visuellen Repräsentationen von abstrakten, nicht physikalisch-basierten Daten zur Wahrnehmungserweiterung zu verstehen. Diese visuell repräsentierten Daten mit Informationsgehalt können erstellt, kalkuliert, oder auf verschiedenste Art und Weisen gefunden werden (z.B. Spielergebnisse einer Sportart, Be-

völkerungsentwicklungsdaten verschiedener Länder, Messergebnisse von wissenschaftlichen Tests). Die Zielsetzung der Informationsvisualisierung ist, von diesen zuvor beschriebenen Daten informiert zu sein oder weitere Informationen und Zusammenhänge durch (visuelle) Analyse zu erhalten (Mazza, 2009).

Informationsvisualisierung ist auch eine Hilfe zur Darstellung und zum Verständnis von abstrakten Daten, welche möglicherweise in keiner Beziehung zum physischen Raum der umgebenden Außenwelt von einem Selbst stehen. Die visuelle Darstellung von Daten hilft in der Kommunikation (z.B. neuen Schreibtisch zusammenbauen und den Plan nur über das Telefon erklärt bekommen, ohne gedruckte Anleitung), in der Analyse von Daten (z.B. Überblick einer Datensammlung mittels eines Diagramms) und beim Bestätigen von Hypothesen (z.B. statistische Ergebnisse als Grafik oder Diagramm) (Mazza, 2009).

### 3.4 Wissenschaftliche Visualisierung

Die wissenschaftliche Visualisierung ist im Gegensatz zur Informationsvisualisierung eine Visualisierung, welcher reale Gegenstände, Strukturen oder einfache Abbildungen zu Grunde liegen, bzw. deren Darstellung noch Ähnlichkeiten mit der zugrundeliegenden Abbildung hat (siehe Abbildung 9).

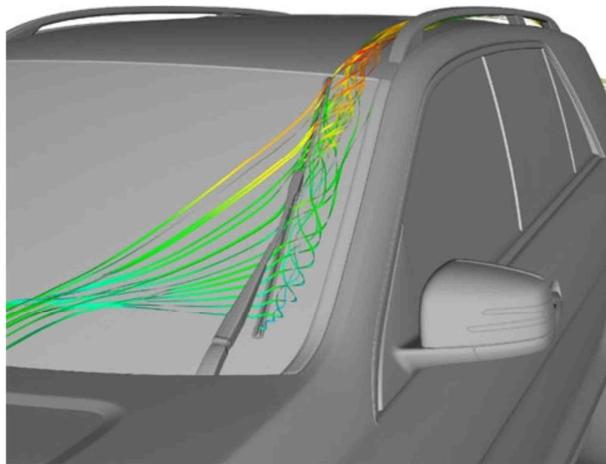


Abbildung 9: Beispiel einer wissenschaftlichen Visualisierung (vgl. Schütz, 2013, p.448)<sup>1</sup>

Der Begriff wissenschaftliche Visualisierung wurde von McCormick, et al. (1987) geprägt und bezeichnet eine Methode, abstrakte Daten auf visuelle Attribute abzubilden. Ziel ist es menschliche visuelle Fähigkeiten zur Erforschung sehr großer Datenmengen zu nutzen, welche aufgrund Ihrer Größe nicht mehr manuell verarbeitet werden können. Durch geeignete graphische Modellierung und Darstellung, sollen der Austausch von wissenschaftli-

---

<sup>1</sup> Luftströmungsvisualisierung durch Stromlinien eines Autoscheibenwischers: Visualisierung des Wirbelsystems hinter dem Scheibenwischer (Schütz, 2013, p.448).

chen Forschungsergebnissen und Arbeiten unterstützt und innere sonst verborgene Zusammenhänge aufgedeckt werden (Schumann, 2004).

Senay und Ignatius (1994) sehen als Ziel der wissenschaftlichen Visualisierung, die interdisziplinäre Unterstützung der Forscher, beim Verständnis von Zusammenhängen, bei Beweisen oder widerlegen von Hypothesen, sowie beim Entdecken neuer Phänomene durch graphische Techniken. Sie definieren das Hauptziel der Datenvisualisierung als: *„eine Erkenntnis aus einem Informationsraum zu gewinnen, indem Daten einfachen graphischen Elementen zugeordnet werden“* (Senay & Ignatius, 1994).

Die zugrundeliegenden Daten von wissenschaftlichen Visualisierungen sind meist sehr groß, dies bedeutet manuell nicht mehr bearbeitbar, wobei die Berücksichtigung der Effizienz und somit die Zeit- und Speicherkomplexität von Algorithmen notwendig ist (Preim & Bartz, 2007). Ein Algorithmus ist ein automatisierbares Verfahren, welches vorgegebene Eingabeinformationen aufgrund eines Systems von Anweisungen, Befehlen, Rechenschritten und so weiter in Ausgabeinformationen umformt (Barth, 2013).

Die wichtigsten Ziele und Forschungsszenarios der wissenschaftlichen Visualisierung sind laut Preim und Bartz (2007):

- Die Erforschung von Daten (nicht gezielte Suche ohne spezifische Hypothese).
- Die Überprüfung einer Hypothese basierend auf Messungen oder Simulationen und deren Visualisierung.
- Die Präsentation von Resultaten.

### **3.5 Medizinische Visualisierung**

Wie unter Punkt 3.2 beschrieben, ist Visualisierung ein Prozess der zum besseren Verständnis von Daten führen soll. Laut den Schriften von McCormick, et al. (1987) über wissenschaftliche Visualisierung, ist der Zweck der Visualisierung *„die Erkenntnis und nicht die Bilder“*. Wird *„Erkenntnis“* als Ziel der Visualisierung angenommen, ist es umso wichtiger vor der Erstellung einer Visualisierung zu verstehen, welche Art von *„Erkenntnissen“* die BenutzerInnen erreichen wollen. Für medizinische Visualisierungssysteme ist ein tieferes Verständnis der Diagnoseprozesse, der Therapieentscheidungen und der intraoperativen Informationsstandards unverzichtbar, um maßgeschneiderte Computerunterstützung anbieten zu können. Ebenso essentiell ist die Berücksichtigung organisatorischer und technischer Restriktionen (Preim & Bartz, 2007).

Die medizinische Visualisierung beschäftigt sich also explizit mit medizinischen Fällen, welchen die generellen Visualisierungsziele zu Grunde liegen (siehe Ende Abschnitt Wissenschaftliche Visualisierung im vorherigen Kapitel 3.4). Ob ein Patient an einer bestimmten Krankheit leidet oder nicht, ist letztlich eine Hypothese welche durch klinische Untersuchungen und medizinischer Bildgebung belegt werden muss. Wenn ein(e) ÄrztIn aufgrund der von PatientInnen beschriebenen Symptome und oder von klinischen Untersuchungen, eine Krankheit nicht erfolgreich feststellen kann, werden radiologische Bilddaten auch ohne explizite Hypothese erstellt. Durch Computerunterstützung vor allem in der Bildbearbeitung, der quantitativen Bildanalyse und der Visualisierung, kann eine radiologische Diagnose aufschlussreich sein (Preim & Bartz, 2007).

Nach der Diagnose von RadiologInnen, in welcher das Stadium und der Grad der Erkrankung festgelegt wird, werden bestimmte Visualisierungen erstellt um die Diagnose den behandelten ÄrztInnen zu präsentieren. Diese Visualisierungen dienen als Aufbereitung der Diagnosedaten und können Messungen (z.B. Darstellung der Position und des Ausmaßes von pathologischen Strukturen) und Kommentare (z.B. eingekreiste Regionen und Pfeile) enthalten, um die Entscheidungsgrundlage hervorzuheben. Die Komplementierung der Untersuchung durch Visualisierungen, soll die Therapieentscheidung unterstützen. Die Ergebnisdaten der Visualisierung der Untersuchungsbefunde können auch zur Diskussion unter den Kollegen verwendet, als Teil von Publikationen fungieren und werden als Unterrichtsunterlagen eingesetzt.

Die Limitierungen der Daten sollte immer berücksichtigt werden. Deren diskrete Eigenschaften definieren die Grenzbedingungen derer Interpretation und Analyse. Zum Beispiel könnten bei spezifischen Strukturen (z.B. Tumore) nicht in Ihrer vollen Größe angezeigt werden oder die gewünschten Strukturen sind in der Realität sehr klein, sodass deren Analyse möglicherweise zu einer hohen Fehlerrate führen würde oder dann hochgradig subjektiv ist (Preim & Bartz, 2007).

In der Humanmedizin wäre das klassische Beispiel für medizinische Visualisierung ein Röntgenbild, aber auch alle bildgebenden Verfahren wie z.B. die **Positronen-Emissions-Tomographie (PET)** oder die **Magnetresonanztomographie (MRT)**. Bei diesen Verfahren kann man zugrundliegende Strukturen noch erkennen, z.B. das Gehirn kann noch als solches erkannt oder erahnt werden, da eine gewisse räumliche Äquivalenz besteht. Bei Blutwerten jedoch gibt es kein bildliches Original, das abgebildet werden könnte. Die Informationsvisualisierung in der Medizin ist somit eine Erweiterung der medizinischen Visualisierung, um in der Humanmedizin die Visualisierung von nicht realen Gegenständen, Strukturen oder zu Grunde liegenden Abbildungen zu beschreiben.

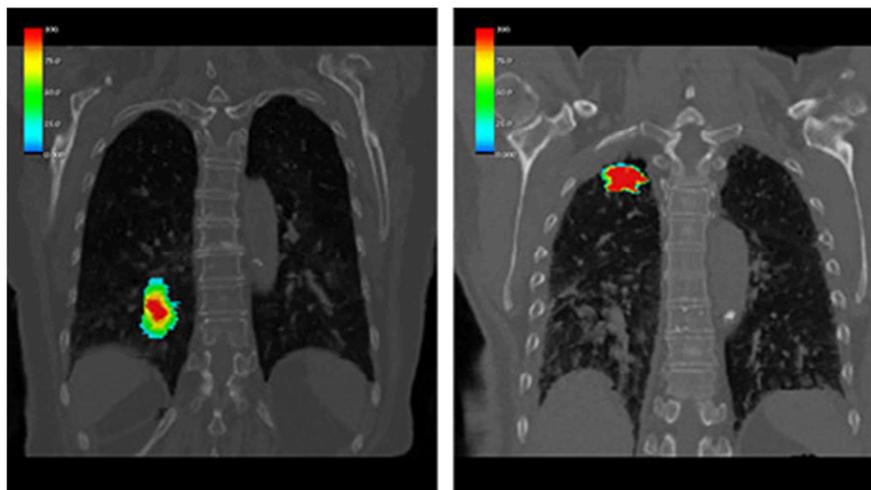


Abbildung 10: Beispiel für medizinische Visualisierung (vgl. Homepage der Universität zu Lübeck, 2014)<sup>2</sup>

### 3.6 Visualisierung von zeitorientierten Daten

In Applikationen werden zeitorientierte Daten dann verwendet, wenn die Entwicklung der Datenwerte über die Zeit von Interesse ist. Einige Beispiele dafür sind Krankenakten, Projektmanagementpläne, Simulationen und Finanzdaten. All diese Beispiele enthalten Informationen, welche über die Zeit beobachtet werden, um daraus Erkenntnisse zu gewinnen. In Aigner, et al. (2007) wurde eine systematischen Kategorisierung von Visualisierungstechniken für zeitorientierter Daten vorgestellt. Es wurde untersucht welchen Einfluss die Darstellung von zeitorientierten Daten auf die analytischen und visuellen Ansätze haben kann. Es wurden drei Hauptkriterien für die Kategorisierung definiert (Aigner, et al., 2007):

1. *time: What are the characteristics of the time axis?* (Zeit: Was sind die Charakteristiken der Zeitachse?)

Unter den Charakteristiken der Zeitachse wurde zwischen Zeitpunkt (*time point*) und Zeitintervall (*time interval*) differenziert, da eine Zeitachse beide darstellen kann. Der wesentliche Unterschied zwischen den beiden Charakteristiken ist, dass bei den Zeitpunkten keine Informationen über die Entwicklung zwischen den Messpunkten vorhanden sind. Dieser Fakt sollte in der Visualisierung berücksichtigt werden, da es sonst zu Fehlinterpretationen kommen kann. Damit der Betrachter sich bei der Darstellung von Zeitintervallen orientieren kann, sollte der Bereich visualisiert werden in denen die intervallbasierten Daten auch Gültigkeit haben. TimeRider verfügt über beide Darstellungsformen, da das Streudiagramm sowohl statisch (=Zeitpunkt), als auch animiert (=Zeitintervall) betrachtet werden kann (Aigner, et al., 2007).

<sup>2</sup> Farbkodierte Visualisierung der Aufenthaltswahrscheinlichkeit von Lungentumoren von zwei Patienten, dargestellt je für ein koronares Schnittbild (Homepage der Universität zu Lübeck, 2014).

## 2. *data: What is analyzed?* (Daten: Was wird analysiert?)

Das zweite Kriterium behandelt Daten, welche mit der Zeitachse verknüpft sind. Es gibt drei Unterkategorien, nach denen Daten eingeteilt werden können. Der erste ist der Bezugsrahmen (*frame of reference*), indem der Kontext der Daten unterschieden wird in abstrakt (*abstract*) oder räumlich (*spatial*). In Aigner, et al. (2007) werden als abstrakten Daten jene bezeichnet, die per se nicht abhängig sind von einer räumlichen Anordnung. Im Gegensatz dazu stehen räumliche Daten, welche eine räumliche Anordnung entweder durch natürliche oder künstliche Umständen enthalten. Der zweite Rahmen ist die Anzahl der zeitabhängigen Variablen (*number of variables*), diese können eindimensional (*univariate*) und somit einfache Datenwerte enthalten oder mehrdimensional (*multivariate*) und somit mehrere Datenwerte enthalten. Der dritte und letzte Bezugsrahmen unter diesem Punkt ist der Grad der Abstraktion (*level of abstraction*) bei einer großen Mengen an Daten (*data*) kann es für den Betrachter von Nutzen sein wenn die Daten abstrahiert (*data abstractions*) und damit übersichtlicher werden (Aigner, et al., 2007).

## 3. *representation: How is it represented?* (Darstellung: Wie wird etwas dargestellt?)

Das letzte Kriterium in Aigner, et al. (2007) beschäftigt sich mit der visuellen Darstellung von zeitorientierten Daten. Es wurden hierfür nur zwei Unterkriterien betrachtet. Die erste Unterkategorie ist die Abhängigkeit der Darstellung von der Zeit (*time dependency*). Es gibt zwei verschiedene Abhängigkeiten, eine ist die statische (*static*) Darstellung von zeitorientierten Daten, in der Bildern sich über die Zeit nicht automatisch verändern. Ein Beispiel für eine statische Visualisierung ist die „*Small Multiples*“ Darstellung, wie in Kapitel 8.2 kurz erwähnt (siehe Exkurs im Anhang C). Die Andere ist die dynamische (*dynamic*) Darstellung, in der die physikalische Dimension „Zeit“ dazu benutzt wird um die Abhängigkeiten der Daten von der Zeit zu vermitteln und in der sich die Darstellung automatisch ändert, wie in Diashows oder Animationen (siehe Kapitel 4). In Neubauer (2012) wird eine Diashow als Darstellung mit sukzessiv wenigen Bilder in diskreten Schritten bezeichnet und die Animation als eine Visualisierung mit einer hohen Bildwiederholungsrate, welche benötigt wird um schnell bewegende Objekte darzustellen.

Die zweite Unterkategorie ist die Räumlichkeit (*dimensionality*), wobei zwischen zweidimensionaler (2D) und dreidimensionaler (3D) Darstellungen unterschieden wird. Laut Aigner, et al. (2007), ist die Verwendung der jeweiligen Darstellung (2D vs. 3D) abhängig von der Aufgabenstellung und von den verfügbaren Daten (Aigner, et al., 2007).

### 3.7 Kurze Abgrenzung von Informationsvisualisierung zu Visual Analytics

In dieser Arbeit dient die Informationsvisualisierung als Grundlage, welche laut Card, et al. (1999) die computergestützte, interaktive, visuelle Repräsentation zum besseren Verständnis von abstrakten Daten ist.

Der verwandte Bereich „*Visual Analytics*“ hingegen, ist ein interdisziplinäres Forschungsgebiet, welches „*Data Mining*“ und Informationsvisualisierung vereint. Hier besteht die Herausforderung, sehr große und multidimensionale Datenmengen zu visualisieren, welche nicht mehr effektiv manuell verarbeitet werden können, um den BenutzerInnen interaktiv nutzbar zu machen (Homepage des Fraunhofer-IGD, 2013). Neben Informationsvisualisierung wird das Forschungsgebiet Data Mining grundsätzlich als eine statistische und systematische Methode bezeichnet, mit dem Ziel neue Muster in einem Datenbestand zu erkennen (Fayyad, et al., 1996).

Die wissenschaftlichen Methoden der Visual Analytics sind bereits für fundamentale Fortschritte in der Genforschung verantwortlich. In der Genforschung müssen Informationen zwischen verschiedenen Molekülen über Gen- und Signalnetzwerke und von der Zelle zum Organismus, sowie von verschiedenen Quellen und Formaten analysiert werden (Wong & Thomas, 2004).

Bei Bedarf werden weitere Technologien bei Visual Analytics herangezogen, z.B. die Analyse von unstrukturierten textbasierten Massenmeldungen. Um diese zu untersuchen, werden zuerst Datensignaturen für die Quelldaten entwickelt, um danach „*high dimensional*“ Repräsentationen zu erzeugen. Dies sind Datenmengen welche zu groß oder zu komplex sind oder sich zu schnell ändern, um sie mit händischen und klassischen Methoden der Datenverarbeitung auszuwerten (Homepage des internationalen Computerbuchverlages O'Reilly Media, 2012). Diese Aufbereitung der Daten wird sowohl statistisch als auch semantisch in verschiedenen Abstraktionslevels und in variablen Kontexten für die Analyse durchgeführt. Das Ziel ist während der Betrachtung des gesamten Kontextes versteckte, schwache oder manchmal auch fehlende Verbindungen und Beziehungen zu entdecken (Wong & Thomas, 2004).

Die Herausforderung in Visual Analytics besteht in der zeitlichen Analyse. Informationsstrukturen, Verbindungen, Rahmenbedingungen und Meinungen verändern sich mit der Zeit und im Kontext der Analyse. Wenn daher Information in eine Analyse eingebunden ist, müssen visuelle und Interaktionsparadigmen zur Verfügung gestellt werden um die Massen an neuen Daten mit digitalen Bibliotheken von Informationen zu Vergleichen, um zu entdecken was sich geändert hat und warum (Wong & Thomas, 2004).

In Visual Analytics ist die Hauptaufgabe der Visualisierung und Interaktionstechniken den BenutzerInnen zu helfen, Erkenntnisse in komplexe Daten und Situationen zu erlangen, wo Modelle alleine ineffizient wären und menschliche Analysefähigkeiten eingesetzt werden sollen. Visualisierung unterstützt nicht immer nur die Repräsentation von kritischen Dateneigenschaften, sondern kann auch kontextbezogene Hinweise bereitstellen, um den BenutzerInnen zu helfen schnell zu interpretieren, was gesehen wird. Des Weiteren ermöglichen Interaktionstechniken den UserInnen über die simple Datenanalyse hinauszugehen und bei der Trend- und Anomalie-Suche, der Auswertung von Hypothesen und der Entdeckung von unerwarteten Zusammenhängen zu helfen (Cook, et al., 2007).

Visual Analytics entstand aus wissenschaftlicher Visualisierung und Informationsvisualisierung, greift jedoch auch auf Methoden und Werkzeuge andere Forschungsfelder bei Bedarf zurück, wie z.B.: statistische Analyse. Der Prozess und die Ziele der Analyse bilden den Ansatz, welcher durch die breite visuelle Schnittstelle des menschlichen Gehirns und des dynamischen Interaktionsstils der Kommunikation erweitert wird (Wong & Thomas, 2004).

## 4 Animation

In diesem Kapitel wird das Thema Animation behandelt, da die Darstellungsform des Programms ein animiertes Streudiagramm ist. In vier Unterkapiteln wird kurz darauf eingegangen was unter Animation im Kontext dieser Arbeit zu verstehen ist.

### 4.1 Allgemein

In Robertson, et al. (2008) wird Animation als eine Veränderung einer visuellen Repräsentation über die Zeit beschreiben, welche die Wahrnehmung von Veränderungen in Visualisierungen über die Zeit erleichtert.

In dieser Magisterarbeit wird der Begriff „*Animation*“ und dessen Verb „*animieren*“, laut der Definition nach dem Duden Online (2013b) verwendet: „[...] *Folge einzelner, den Bewegungsablauf wiedergebender Bilder [...]*“. Damit sind Bilder gemeint, welche über die Zeit sequenziell (aufeinanderfolgend) betrachtet werden und somit die Illusion der Bewegung entsteht. Animation verwendet somit den Parameter „Zeit“ um Zeit darzustellen. Animation ist also eine Sequenz von statischen Repräsentationen (Bildern) für jeden Zeitschritt der Daten, um die Informationen in einer logischen und vordefinierten Reihenfolge darzustellen (Aigner, et al., 2011 und Rob, et al., 1997 zitiert in Kriglstein, et al., 2014). Weitere Aspekte sind Anzeigedatum, Dauer, Änderungsrate, Frequenz und Synchronisation (MacEachren, 2004 zitiert in Kriglstein, et al., 2014). Für einen flüssigen Ablauf einer Animation sind 24 bis 30 Bilder pro Sekunde (24-30 fps = „*frames per second*“) notwendig, für eine Diashow sind jedoch 2-4 fps ausreichend (Aigner, et al., 2011 und Harrower, et al., 2008 zitiert in Kriglstein, et al., 2014).

Die Verwendung von Animation hat Vor- und Nachteile. Einerseits ermöglicht es die Darstellung von komplexen Daten auf einem Bildschirm, da es nicht notwendig ist „Zeit“ als Dimension zu inkludieren. Andererseits können die Betrachter ein Phänomen auf einem Blick sehen, da die kognitive Last des Kurzzeitgedächtnisses drastisch erhöht wird. Daher wiederholen die UserInnen sehr oft eine Animation immer wieder, um die dargestellten Daten zu verstehen (Pohl, et al., 2010).

In Kriglstein, et al. (2012) wird beschrieben, dass eine große Anzahl von Studien in der kognitiven Psychologie über die Wahrnehmung von Bewegung und Veränderungen existiert, welche alle relevant sind für die Verwendung von Animationstechniken um zeitabhängige Daten darzustellen. In der Studie von Rensink, et al. (1997 zitiert in Kriglstein, et al., 2014) wird das Phänomen von Veränderungsblindheit beschrieben und die Fähigkeit von Menschen sich nur auf ein beschränktes Sichtfeld konzentrieren zu können. Des Weiteren beschreibt Pylyshyn (2007 zitiert in Kriglstein, et al., 2014), dass der Mensch nur eine beschränkte Anzahl, durchschnittlich vier gleiche Objekte, von beweglichen Zielen folgen kann (Kriglstein, et al., 2014).

Die Anwendung von Animation wird in der Literatur kontrovers diskutiert. Meistens wird darauf hingewiesen, dass Animation mit Bedacht eingesetzt werden sollte und dass Animation nur dann eingesetzt werden sollte wenn eine statische Darstellung nicht möglich ist (siehe auch Harrower & Fabrikant, 2008 zitiert in Kriglstein, et al., 2014). Der Einsatz von Animationstechniken ist nur unter bestimmten Bedingungen von Vorteil, wie z.B. als Hilfestellung zum Erkennen von Änderungen der Daten und zum besseren Verständnis wie die Daten sich über die Zeit entwickeln bei Aufgabenstellungen mit dynamischer Datenentwicklungen (z.B. Archambault, et al., 2011; Farrugia & Quigley, 2011; McGrath & Blythe, 2004; Nakakoij, et al., 2001; Robertson, et al., 2008; Tekušová & Kohlhammer, 2007; Tekušová & Kohlhammer, 2008; Tekušová & Schreck, 2008 zitiert in Kriglstein, et al., 2014). Der Grund dafür ist wahrscheinlich, dass Animationstechniken kleinste Schritte (*micro steps*) von Datenänderungen anzeigen können im Gegensatz zur Analyse von zeitabhängigen Daten durch multiple kleine statische Grafiken (Slocum, et al., 2001 zitiert in Kriglstein, et al., 2014). McGrath & Blythe (2004 zitiert in Kriglstein, et al., 2014) weisen außerdem daraufhin, dass die Verwendung von Animation in Kombination mit räumlicher Darstellung (z.B. Graphen) eine gute Verdeutlichung der verschiedenen Dateneigenschaften ist. Weitere Studien, wie Zaman, et al. (2011 zitiert in Kriglstein, et al., 2014) beschreiben eine gut designte Animation als machtvolle Technik, die aber nicht jedem Vorteile bringt und sogar verwirrend sein kann. Eine Möglichkeit dafür könnte eben die beschränkte Fähigkeit des Menschen sein, sich auf nur einen kleinen Teil der Information konzentrieren zu können und dadurch kann es passieren das die UserInnen wichtige Informationen für die weitere Analyse nicht wahrnehmen oder bemerken (Kriglstein, et al., 2014).

Für eine effektive Nutzung von Animationstechniken als zusätzliche Hilfestellung ist auf die geeignete Animationsgeschwindigkeit, die Interaktivität, die Verwendung von zusätzlichen Interaktionstechniken und auf den Ablauf der Animation in Relation zu Ordnung und Synchronizität (MacEachren, 2004 zitiert in Kriglstein, et al., 2014) zu achten. Dadurch sollen Probleme wie Verschwinden des Beobachteten, Aufmerksamkeit, Vertrauen und Komplexität gelöst werden. In Midtbø, et al. (2007 zitiert in Kriglstein, et al., 2014) wird beschrieben, dass die Verwendung verschiedener Animationstechniken von der Zeitspanne

abhängig sind. Die ProbandInnen der Kriglstein, et al. (2014) Studie bevorzugten, sowohl zirkuläre Animationen der Zeitspannen, sodass die Muster wiederholt abliefen, als auch lineare Animationen für Zeitspannen ohne zirkulären Charakter (Kriglstein, et al., 2014).

#### 4.1.1 Animationsgeschwindigkeit

In Kriglstein, et al. (2012) wird der Einfluss der Geschwindigkeit der Animation auf die Analyse von zeitabhängigen Daten beschrieben, z.B. um Veränderungen in der Animation zu entdecken. Daher wird empfohlen, dass die Geschwindigkeit der Animation angemessen langsam sein sollte, damit Bewegungen und Veränderungen einfach wahrgenommen werden können (Ware, 2012 zitiert in Kriglstein, et al., 2014). Die optimale Geschwindigkeit für eine Animation zu finden ist eine Herausforderung, da z.B. bei einer zu schnellen Animation, die UserInnen keine Veränderungen der Daten erkennen können und somit Informationen nicht gesehen oder verstanden werden könnten (siehe hierfür Griffin, et al., 2006 und Tekušová, et al., 2008 zitiert in Kriglstein, et al., 2014). Sollte jedoch die Animation zu langsam sein, können die Betrachter eventuell weder Trends noch Gruppierungen der Datenpunkte erkennen, da z.B. die BenutzerInnen die Bewegung der Animation nicht erkennen könnten (Kriglstein, et al., 2014).

Die Ergebnisse von Griffin, et al. (2006 und Tekušová, et al., 2008 zitiert in Kriglstein, et al., 2014), Harrower & Fabrikant (2008 zitiert in Kriglstein, et al., 2014) und Kjellin, et al. (2008 zitiert in Kriglstein, et al., 2014) weisen darauf hin, dass die Geschwindigkeit der Animation abhängig von verschiedenen Aspekten ist, wie die Distanz der sich bewegenden Objekte, die Dauer der Animation und die Bildwiederholungsrate (auch bekannt als Bildfrequenz) der Animation. Kjellin, et al. (2008 zitiert in Kriglstein, et al., 2014) beobachteten ebenfalls das Phänomen der Veränderungsblindheit, die ebenfalls ein wichtiger Faktor für die Wahl der richtigen Geschwindigkeit der Animation ist. Bei Veränderungsblindheit haben die ProbandInnen Schwierigkeiten, große Veränderungen zu erkennen, weil es zu einer visuellen Unterbrechung kam, z.B. weil ein Blicksprung (Augenbewegung) durchgeführt wurde (Kriglstein, et al., 2014).

Daher wird eine Standardgeschwindigkeit für die Animations empfohlen, welche die UserInnen einfach während der Animation interaktiv nach eigenem Ermessen ändern und kontrollieren können, damit mehr Zeit für die Bereiche von Interesse bleibt (Müller & Schumann, 2003 zitiert in Kriglstein, et al., 2014). Weitere Studien zeigen, dass die Möglichkeit die Geschwindigkeit der Animation zu kontrollieren eine Hilfestellung für den Denkprozess (mentales Modell) der BenutzerInnen ist und dass diese schneller frustriert waren, wenn sie keine Kontrolle über die Animationsgeschwindigkeit hatten (Harrower, 2003; Nakakoji, et al., 2001 und Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014).

### 4.1.2 Interaktivität

Damit die Fragen der BenutzerInnen, welche während der Beobachtung der animierten Daten entstehen, beantwortet werden können, reicht es meist nicht nur die Daten durch Beobachtung zu analysieren (Rob, et al., 1997 zitiert in Kriglstein, et al., 2014). Interaktionsmöglichkeiten erlauben es dem Betrachter die Kontrolle über die Animation zu erhalten und dabei die visuelle Untersuchung zu verbessern, im Gegensatz zu deren passiver Betrachtung (Bloc, 2000 und Tversky, et al., 2002 zitiert in Kriglstein, et al., 2014). Die meisten Arbeiten (siehe z.B.: Archambault, et al., 2011; Bezerianos, et al., 2006; Craig, et al., 2005; Farrugia & Quigley, 2011; Midtbø, et al., 2007; Nakakoji, et al., 2001; Rind, et al., 2011; Robertson, et al., 2008; Tekušová & Kohlhammer, 2007 zitiert in Kriglstein, et al., 2014) verwenden den allgemein gültigen und bekannten Standard von VCR-Kontrollelementen (*Video Cassette Recording*), wie Start (*Play*), Pause, Schnellvorlauf (*FF= fast forwarding*) und Schnellrücklauf (*RW= fast rewind*). Die VCR-Funktionen werden in Kombination mit Schiebereglern verwendet, um die Geschwindigkeit einfach kontrollieren zu können, siehe ein Beispiel hierfür in TimeRider in Kapitel 8.2 (Kriglstein, et al., 2014).

UserInnen-Aktionen wurden in einigen Studien analysiert (siehe z.B.: Lowe, 2008; Nakakoji, et al., 2001; Tversky, et al., 2002 zitiert in Kriglstein, et al., 2014) mit dem Ergebnis, dass (i) Interaktivität dabei helfen kann die Schwierigkeiten bei der Wahrnehmung der Daten zu lösen und (ii) das Verstehen von Animationen hilft den UserInnen ein Gefühl für die Daten zu bekommen. Darüber hinaus analysierten Studien (siehe z.B.: Archambault, et al., 2011; Bezerianos, et al., 2006; Boyandin, et al., 2012; Farrugia & Quigley, 2011; Nakakoji, et al., 2001; Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014) in Ihrer Auswertung die Verwendung von Interaktionselementen, während die UserInnen den Datensatz analysierten. Es wurde zum Beispiel beobachtet, dass die BenutzerInnen es vorzogen als Erstes die gesamte Animation zu untersuchen um eine Übersicht zu erlangen. Der nächste Schritt war durch das Verwenden der Kontrollelemente, zu verschiedenen Positionen von Interesse zu springen und um diese detailliert zu analysieren, z.B. für den Vergleich von verschiedenen Datenzuständen. Durch die zusätzliche Interaktionsmöglichkeit, benötigten die TeilnehmerInnen mehr Schritte für die Analyse und mehr Zeit mit der Animation als mit einer statischen Darstellung (Chevalier, et al., 2010; Farrugia & Quigley, 2011; Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014). In Boyandin, et al. (2012 zitiert in Kriglstein, et al., 2014) interagierten die TeilnehmerInnen länger mit der Animationstechnik, als mit der ebenfalls verwendeten *Small Multiples* Visualisierung (siehe hierfür Anhang C) für dieselbe Anzahl von Entdeckungen. Daher wird empfohlen, dass Animationstechniken dann eingesetzt werden, wenn Präzision wichtiger ist als eine schnelle Reaktionszeit (Kriglstein, et al., 2014).

Batram (1997 zitiert in Kriglstein, et al., 2014) und Ware (2012 zitiert in Kriglstein, et al., 2014) beschrieben in Ihren Untersuchungen zur kognitiven Psychologie mit Fokus auf die Wahrnehmung von Bewegung und Veränderung, dass die BenutzerInnen nur vier bis fünf verschiedenen Objekten zur selben Zeit folgen konnten. Dies wurde als mögliche Ursache identifiziert, dass ProbandInnen sich in der Animation verlieren, da die Anzahl der zu beobachtenden Objekte zu groß war und deshalb die Aufmerksamkeit abgelenkt wurde (vgl. Farrugia & Quigley, 2011; Nakakoji, et al., 2001; Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014). Daher sind weitere Maßnahmen zusätzlich zu den Interaktionsoptionen notwendig, wie auswählen (markieren), filtern und Suchfunktionen, um die Komplexität und Informationsüberladung zu reduzieren (Harrower, 2003; Tekušová, et al., 2008 zitiert in Kriglstein, et al., 2014). Ein Beispiel dafür wäre eine Filterfunktion, welche eine spezifische Animationsauswahl ermöglicht (Farrugia & Quigley, 2011; Nakakoji, et al., 2001; Rensink, 2002 zitiert in Kriglstein, et al., 2014). In Tekušová & Kohlhammer (2007 zitiert in Kriglstein, et al., 2014) wurde beobachtet, dass zoomen oder eine „*Details on Demand*“<sup>3</sup> Funktion hilfreich ist, um die Anzahl der angezeigten Informationen zu kontrollieren und zur selben Zeit die Möglichkeit zu geben, mehr oder weniger Informationen der beobachtbaren Daten anzuzeigen oder auszublenden (Kriglstein, et al., 2014).

### 4.1.3 Zusätzliche Optionen

Zusätzlich zu den Interaktionstechniken, welche den Ablauf der Animation steuern können, wird eine Kombination aus weiteren Optionen, wie Farbänderung, Spurenanzeige oder Klang, in den Studien empfohlen (siehe auch z.B. Bezerianos, et al., 2006; Chevalier, et al., 2010; McGrath & Blythe, 2004; Midtbø, et al., 2007; Rind, et al., 2011; Schlienger, et al., 2007; Tekušová & Kohlhammer, 2007; Zaman, et al., 2011 zitiert in Kriglstein, et al., 2014). Weitere Zeichen wie Farben oder Klänge können dazu verwendet werden, die BenutzerInnen auf spezifische Bereiche in der Animation oder auf spezifische animierte Objekte hinzuweisen. Die Spurendarstellung (siehe ein Beispiel aus TimeRider in Kapitel 5.6) kann dazu verwendet werden die historische Entwicklung der Datenpunkte anzuzeigen und somit die BenutzerInnen bei der Betrachtung der Datenveränderungen über die Zeit zu unterstützen. Einerseits wurde in Studien beobachtet, dass die Verwendung der Spurendarstellung den BenutzerInnen dabei half die Trends der Daten schneller zu erkennen, als ohne eine solche Funktion (siehe z.B. Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014), aber andererseits zeigten Studien (vgl. Rind, et al., 2011; Robertson, et al., 2008; Tekušová & Kohlhammer, 2007 zitiert in Kriglstein, et al., 2014), dass diese zusätzliche Option nach Aktivierung die UserInnen verwirrte, vor allem bei großen Datensätzen mit vielen Datenpunkten oder bei mehreren Datensätzen gleichzeitig (Kriglstein, et al., 2014).

---

<sup>3</sup> „The *details on demand*-technique allows interactively selecting parts of data to be visualized more detailed while providing an overview of the whole informational concept.“ - (Homepage der InfoVis-Wiki Seite, 2006)

#### 4.1.4 Animation vs. statische Darstellung

Der Vergleich zwischen Animationsvisualisierungen und einer statischen Darstellung wie *Small Multiples* (siehe hierfür Anhang C) und welche Darstellung effektiver ist, hängt von der Aufgabenstellung ab. Für Präsentationszwecke wird eine Verwendung von Animation bei bestimmten Datenperspektiven und deren Veränderung über die Zeit (z.B. um Datentrends zu erkennen) empfohlen. Meistens wird eine korrespondierende statische Darstellung als zu überladen wahrgenommen, vor allem bei *Small Multiples* Darstellungen wurden die zeitabhängigen Daten als zu klein und als zu unübersichtlich empfunden (siehe z.B. Boyandin, et al., 2012; Chevalier, et al., 2010; Fabrikant, et al., 2008; Nakakoji, et al., 2001; Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014). Des Weiteren war es für die TeilnehmerInnen bei der Animation einfacher, den Änderungen der Daten über die Zeit durch die bewegten Bilder zu folgen, als bei einem Vergleich mehrere statischer Bilder zu verschiedenen Zeitpunkten (vgl. McGrath & Blythe, 2004; Tekušová & Kohlhammer, 2007 zitiert in Kriglstein, et al., 2014) (Kriglstein, et al., 2014).

Die Verwendung von Animation ist auch zur Identifizierung von Mustern hilfreich. In Griffin, et al. (2006 zitiert in Kriglstein, et al., 2014) wird beschrieben, dass die Erkennung von Mustern davon abhängig ist, wie sichtbar die sich bewegenden Muster waren. Im Vergleich zu einer *Small Multiples* Darstellung zeigte sich, dass die TeilnehmerInnen durch Animation in derselben Zeit mehr Muster entdeckten. In Boyandin, et al. (2012 zitiert in Kriglstein, et al., 2014) wurde dies bestätigt und darauf hingewiesen, dass durch die Verwendung von Animation, lokale Muster und plötzliche Änderungen effektiv identifizieren werden könnten (Kriglstein, et al., 2014).

Für Analyseaufgaben zeigten Studien (Archambault, et al., 2011; Griffin, et al., 2006; Robertson, et al., 2008 zitiert in Kriglstein, et al., 2014) jedoch, dass statische Darstellungen als effektiver als Animationen empfunden wurden, z.B. für die Analyse von topologischen Informationen oder um Informationen an einem bestimmten Zeitpunkt zu vergleichen. Obwohl kein signifikanter Unterschied in der Leistung beobachtet wurde, wiesen Kjellin, et al. (2008 zitiert in Kriglstein, et al., 2014) daraufhin, dass der Lerneffekt der ProbandInnen mit statischer Darstellung höher war, als mit Animationsvisualisierung. Ebenfalls hoben sie hervor, dass zu diesem Effekt jedoch weitere Untersuchungen durchgeführt werden sollten (Kriglstein, et al., 2014).

#### 4.1.5 Größe der Datensätze

Die Größe der verwendeten Datensätze spielt laut Robertson, et al. (2008 zitiert in Kriglstein, et al., 2014) eine wichtige Rolle, abhängig ob die Größe die ProbandInnen in der Lösung derer Aufgaben unterstütze. Es wurde beobachtet, dass obwohl die Testpersonen

die Verwendung der Animation als „lustiger“ und interessanter als die statische Darstellung empfanden, waren sie nur bei kleinen Datensätzen durch die Animation erfolgreicher und bei größeren Datensätzen durch die statische Darstellung (Kriglstein, et al., 2014).

Am Ende wiesen Kriglstein, et al. (2014) jedoch darauf hin, dass es noch zu wenig Forschung im Bereich Animationsvisualisierungen und deren Nach- und Vorteile gibt, um generelle Richtlinien definieren zu können. Zukünftiger Forschung obliegt es umfangreichere Rahmenbedingungen für Richtlinien in diesem Forschungsbereich zu erarbeiten (Kriglstein, et al., 2014).

## 4.2 Animation in der Trendanalyse

Animation wurde bei Benutzeroberflächen (*user interface*) für verschiedene Zwecke verwendet (Baecker & Small, 1990 zitiert in Robertson, et al., 2008). Animationen wurden in Visualisierung verwendet um verschiedenste Veränderungen zu zeigen: wie Datentransformationen in verschiedene Zustände (Robertson, et al., 1991; Robertson, et al., 2002; Heer & Robertson, 2007 zitiert in Robertson, et al., 2008), Übergänge zwischen verschiedenen Darstellungen (Bederson & Boltman, 1999 zitiert in Robertson, et al., 2008), Veranschaulichung wie etwas funktioniert (Zongker & Salesin, 2003 zitiert in Robertson, et al., 2008) und Analyse von Trends (*Gapminder, MicroStrategy Animated Bubble Chart, Report Portal Moving Bubble Chart* zitiert in Robertson, et al., 2008).

Die Studie von Kehoe, et al. (2001 zitiert in Robertson, et al., 2008) demonstriert, dass die Verwendung von Animation für die Trendanalyse wertvoll sein kann. Animation kann zu einem besseren Verständnis führen und somit die Motivation verstärken und ein komplexes Thema zugänglicher machen. Eine der ersten Verwendungen von Animation in Visualisierungen war die Animation von Algorithmen (Robertson, et al., 2008).

Batram (1998 zitiert in Robertson, et al., 2008) behauptet: „*Animation ruft die Eigenheiten von Gruppierungen hervor, welche quer in einer dichten Datendarstellung vieler, gleicher Bewegungen stattfinden*“. Dadurch ermöglicht Animation den UserInnen auf Anhieb verbundene Elemente zu erkennen, welche unter anderem auch weit verstreut sein können. Somit können Trends effektiv mit Animationen dargestellt werden. Bewegen sich Gruppen von Datenelementen alle generell in dieselbe Richtung, wird ein genereller Trend erkannt und auch ein Gegentrend kann wahrgenommen werden. Die in Ware (2004 zitiert in Robertson, et al., 2008) beschriebene Beobachtung, dass das menschliche Gehirn eine starke Tendenz hat, bewegende Objekte hierarchisch zu gruppieren, unterstützt die Hypothese der Trenderkennung durch das menschliche Gehirn (Robertson, et al., 2008).

Im Gegensatz zur Übergangs/Wechsel-Animation, welche in erster Linie den BenutzerInnen helfen den Zusammenhang zu verstehen, vermittelt Trendanimation eine Meinung. Die Zeiteinteilung (*timing*) ist länger, da die Komplexität der Bewegung und die Anomalien von Interesse, nur in der Bewegung gesehen werden. Die Übergangs/Wechsel Animation entsteht zwischen dem Wechsel von einem Standbild zum nächsten neuem Standbild, wohingegen die Trendanimation sich kontinuierlich zwischen den Zuständen der Daten bewegen muss. Die optimale Zeit für eine Trendanimation ist nicht bekannt. Diese Frage wird teilweise durch die Interaktion gelöst, indem die BenutzerInnen die Kontrolle über die Abspielgeschwindigkeit besitzen und somit je nach Belieben pausieren und auch Vor- oder Rücklauf betätigen können (Robertson, et al., 2008).

Cleveland & McGill (1985 zitiert in Robertson, et al., 2008) beschreiben wie einfach verschiedene grafische Eigenschaften wahrgenommen und beurteilt werden können. Deren Ergebnisse zeigen, dass Datenpositionen weit genauer beurteilt werden als eine Datenfläche (Robertson, et al., 2008).

Tversky, et al. (2002 zitiert in Robertson, et al., 2008) sind jedoch skeptisch gegenüber vielen veröffentlichten Animationsstudien und zeigen die Probleme der Untersuchungsmethode Animation auf. In deren eigenen Studien in denen sie Animation zum Aufzeigen oder Darstellen der Funktion komplexer Systeme verwendeten, konnten Sie keinen Vorteil darin finden. Nur die Möglichkeit eines Vorteils für Animationen bei Übergängen oder Wechseln wird erwähnt. Die Effektivität von Animationen für Trends wird jedoch bezweifelt (Robertson, et al., 2008). Zongker & Salesin (2003 zitiert in Robertson, et al., 2008) widersprechen Tverskys Behauptung, in dem Sie einige Designgrundlagen aufzählen, welche Präsentationen mittels Animation als effektiv belegen (Robertson, et al., 2008).

Griffin, et al. (2006 zitiert in Robertson, et al., 2008) vergleichen animierte Abbildungen mit statischen kleinen multiplen Darstellungen und beschreiben die signifikanten Vorteile sowohl der Geschwindigkeit, als auch der Exaktheit der animierten Abbildungen (Robertson, et al., 2008).

### **4.3 Animation in TimeRider**

In TimeRider werden Animationen als Visualisierung verwendet (für Details siehe vorheriges Kapitel 4.1), daher werden in diesem Kapitel verwandte Arbeiten und deren Sichtweise darauf kurz vorgestellt. Mit dem Fokus auf Animation als zusätzliche Darstellungsdimension, welche die Repräsentation von großen multivariaten dynamischen Daten vereinfacht, werden in Batram (Batram, 1997 zitiert in Rind, et al., 2011) potenzielle Vorteile und Anwendungen von Animationen beschrieben. Im Detail werden fünf Vorteile der

Verwendung von Bewegungsabläufen genannt: (i) Wahrnehmungseffizienz, (ii) mehrere Interpretationsmöglichkeiten, (iii) noch nicht häufig in Verwendung, (iv) mehr potentielle Codierungsbandbreite, (v) technologisch immer einfacher umsetzbar. In Nakakoji, et al. (2001 zitiert in Rind, et al., 2011) wurde anhand von zwei Benutzerstudien die kognitiven Effekte und die Interaktivität von animierten Visualisierungen für Forschungsanalysen untersucht. Die Ergebnisse unterstützen die Argumente von Bartram (Bartram, 1997 zitiert in Rind, et al., 2011) und beschreiben Animation als ein mächtiges Instrument. In Griffin, et al. (Griffin, et al., 2006 zitiert in Rind, et al., 2011) wird die Wirksamkeit von Animation auf das Erkennen von bewegenden Mustern und Clustern in einer 2D Darstellung untersucht, unter Berücksichtigung der Effekte von zeitlicher Abstimmung (z.B. Bildfrequenz). Die Effektivität wurde gemessen und mit der *Small Multiples* Darstellung verglichen. Dies ist eine Visualisierung einer Serie von Graphen oder Tabellen, bei gleicher Skalierung und Achsen, damit die Entwicklung von verschiedenen Zeitpunkten oder Objekten einfach verglichen werden kann (siehe Abbildung 11 und für eine detaillierte Beschreibung Anhang C). Sowohl bei der Fehlerrate als auch der verstrichenen Zeit um eine Aufgabe zu erfüllen, wurde ein Vorteil im Vergleich zur *Small Multiples* Darstellung für Animation beobachtet (Rind, et al., 2011).

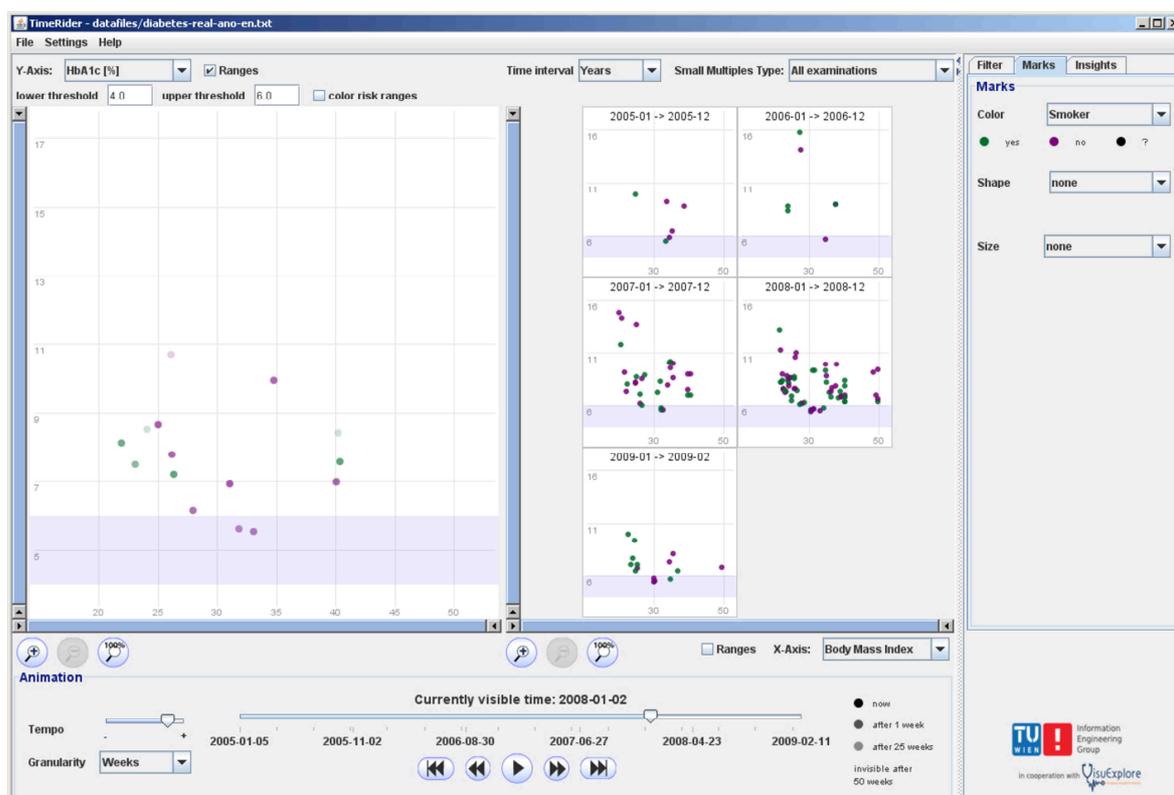


Abbildung 11: Small Multiples Darstellung in einer weiterentwickelten Version von TimeRider (Neubauer, 2012, p.46).

## 4.4 Animated Scatter Plot

„*Animated Scatter Plot*“ heißt zu Deutsch „*animiertes Streudiagramm*“. Dieses wird in diesem Unterkapitel beschrieben, da dies die Darstellungsform von TimeRider ist. Wobei „*animieren*“ laut dem Duden Online (2013b) bedeutet: „[...] *Folge einzelner, den Bewegungsablauf wiedergebender Bilder* [...]“.

Streudiagramme (oder Punktdiagramme) sind hingegen eine visuelle Methode, um einen Zusammenhang von zwei Variablen auf einem kartesischen Koordinatensystem darzustellen. Die Ausprägung der einen Variable wird z.B. auf der X-Achse eingetragen, die Ausprägung der anderen Variable wird z.B. auf der Y-Achse aufgetragen siehe (Homepage des „KSA“, 2013).

Durch das Auftragen der Werte der Variablen auf den Achsen erhält man eine Reihe von Schnittpunkten. Das sich aus den Schnittpunkten ergebende Diagramm wird auch Verteilungsgrafik (auch Streuungsdiagramm, Streudiagramm oder Scatter Plot) genannt, siehe Abbildung 12 unten (Homepage des „KSA“, 2013).

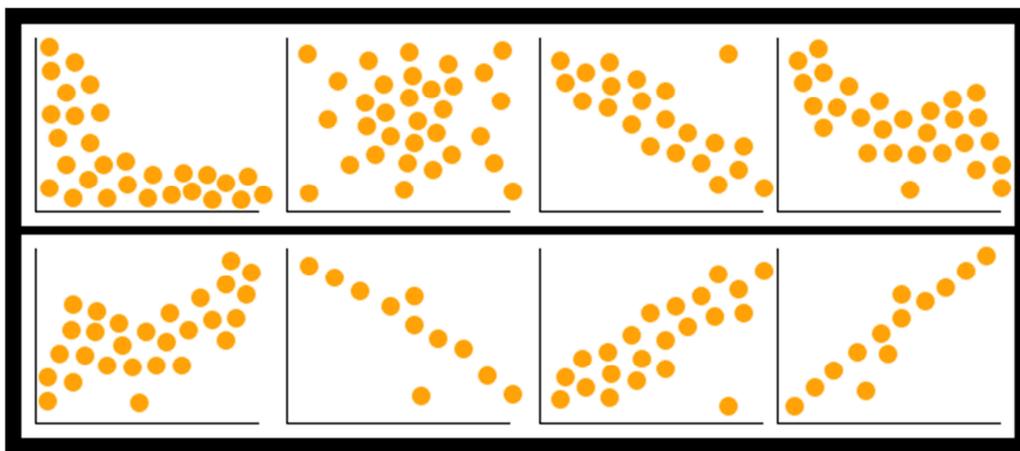


Abbildung 12: Beispielbilder von einigen zufällig gewählten Streudiagrammen (in Anlehnung an Homepage des „KSA“, 2013).

Streudiagramme eignen sich zur grafischen Darstellung bivariater Daten, also zur Darstellung von Wertepaaren in einem Koordinatensystem. An Lage und Dichte des „*Datenpunkteschwarms*“ lässt sich ablesen, ob ein Zusammenhang zwischen den Variablen besteht. Wobei „*Schwarm*“ in diesem Fall, als größere Anzahl gleichartiger Schnittpunkte zu verstehen ist (Homepage des „Duden“, 2013c). Sie werden folgendermaßen erstellt: Man trägt in einem Diagramm auf den gewünschten Achsen die Schnittpunkte der beiden Variablen ein, dort wo beide Einträge aufeinander treffen, wird eine Markierung (ein Punkt) eingefügt. Das Streudiagramm eignet sich somit zum Aufzeigen des Zusammenhangs zwischen Variablen, der Darstellung der Regression (= Beschreibung der Beziehung oder Abhängigkeiten von Variablen untereinander, siehe Homepage des „Duden“, 2013d) und der Korrelation (= statistischer Zusammenhang, siehe Homepage des „Duden“, 2013e) (Homepage

des „KSA“, 2013). Dem Betrachter ist somit die Möglichkeit gegeben einen statischen Auszug eines eventuell komplexen Sachverhalts auf einen Blick grafisch zu betrachten, wie eine Fotografie. Anhand des Beispiel von TimeRider in Abbildung 13, lässt sich ablesen, dass mit höherem *BMI* (=Body-Mass-Index auf der X-Achse) auch ein erhöhter *CHOL-Wert* (=Gesamtcholesterin auf der Y-Achse) im Patientenkollektiv gemessen wurde.

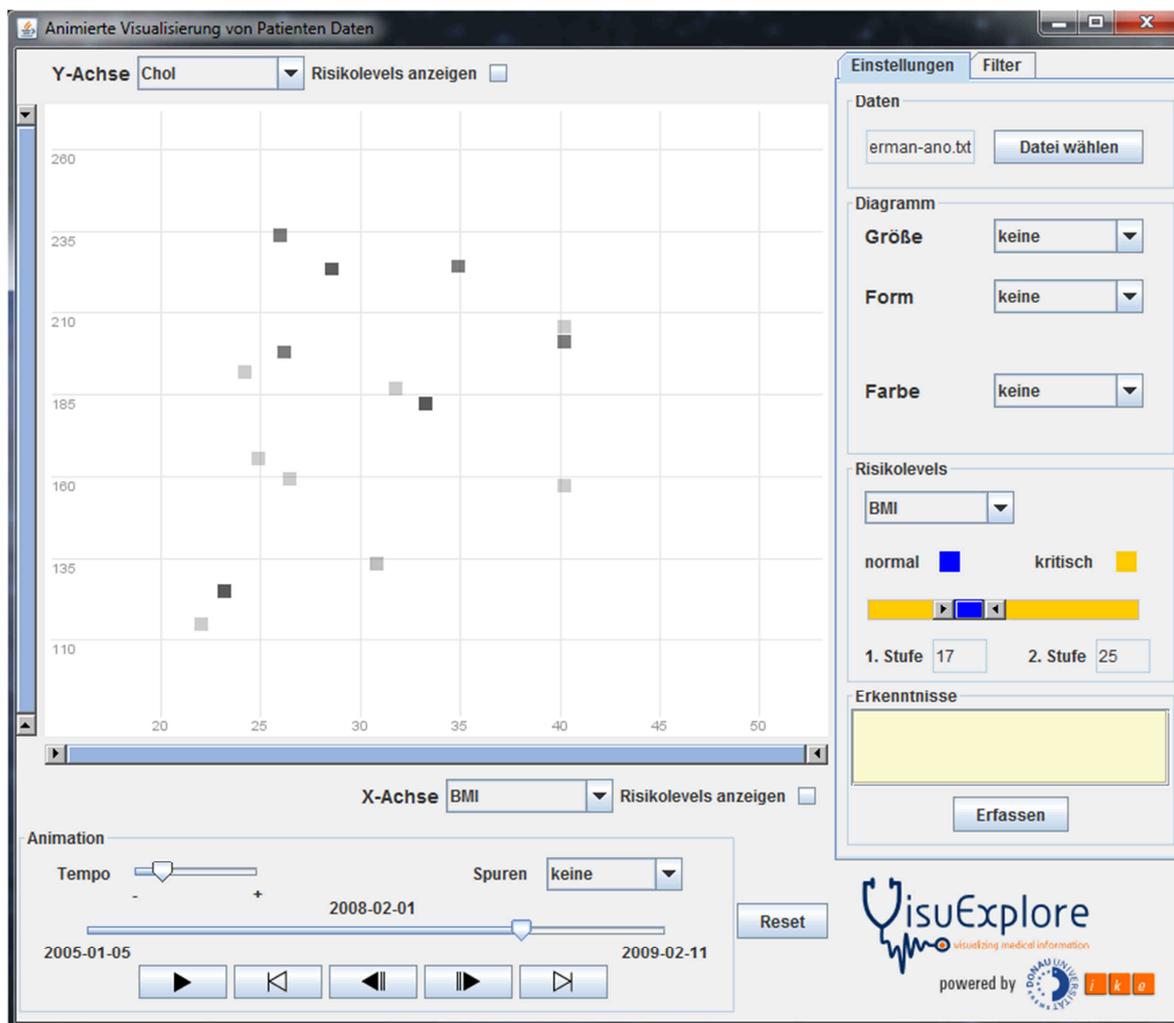


Abbildung 13: Datenpunkteschwarm in TimeRider.

Damit diese statische Darstellung in eine dynamische über die Zeit umgewandelt werden kann, können zwei Möglichkeiten verwendet werden. Eine davon ist die Verwendung der räumlichen Funktion einer Zeitachse bei einer statischen Repräsentation oder eine dynamischen Darstellung, welche verschiedene Zeitpunkte nach und nach anzeigt (z.B.: eine Animation, wie in TimeRider implementiert, siehe hierfür Abbildung 14 oder eine Diavorführung) (Aigner, et al., 2007 zitiert in Rind, et al., 2011). Da für die Darstellung von Scatter Plots beide Achsen für Datenvariable verwendet werden, wird die Animation zur Darstellung von Entwicklungen über die Zeit verwendet.

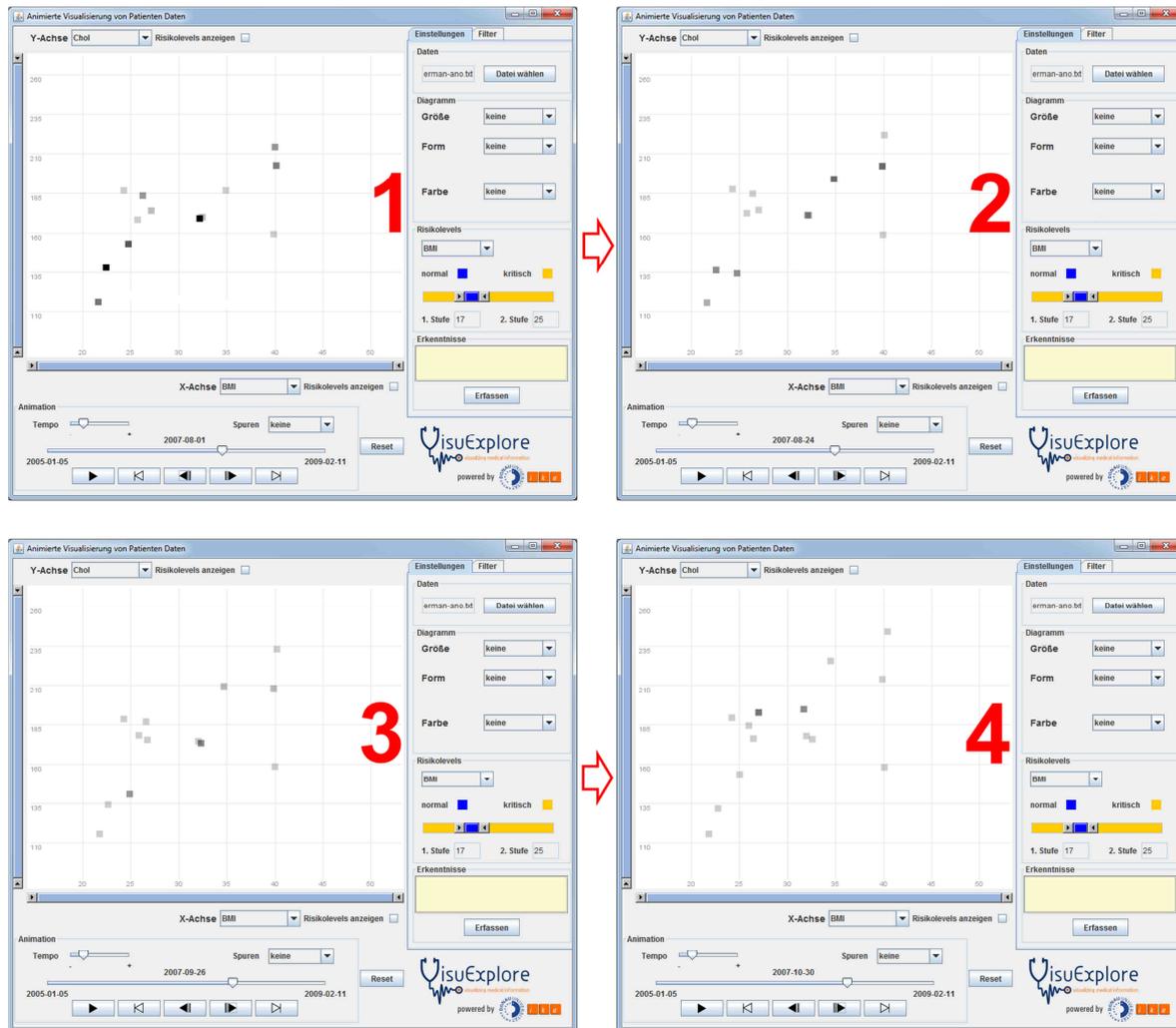


Abbildung 14: Animation in TimeRider<sup>4</sup>

Animated Scatter Plot ist somit ein Streudiagramm, welches animiert (bewegt) dargestellt wird. Dadurch kann das Verhalten des „Punkteschwarm“ in verschiedenen Zuständen analysiert werden, indem z.B. verschiedene Schnittpunktewerte über die Zeit betrachtet werden. Die Punktwolke des Schwarms kann sich dadurch verändern und „bewegen“ oder eben auch nicht. Je nach entstehender Veränderung, ob Animation oder Stillstand, können dann die BenutzerInnen aus dem entsprechenden Kontext heraus, neue Erkenntnisse und Einsichten, in dem zu untersuchenden Sachverhalt, gewinnen. Animated Scatter Plot wurden in Rosling (Rosling, 2007 zitiert in Rind, et al., 2011) als eine effektive Darstellung von komplexen sozioökonomischen Beziehungen über die Zeit verwendet. Ein reales Beispiel wäre hierfür eine medizinische Diagnosefindung, wo die zuständigen ÄrztInnen durch eine Visualisierung mittels eines animierten Streudiagramms in deren Annahme bestätigt oder nicht bestätigt werden.

<sup>4</sup> Die Abbildung zeigt einen Ausschnitt von vier verschiedenen Messzeitpunkten, wobei die roten Zahlen zur Orientierung nachträglich vom Autor dieser Arbeit eingefügt wurden. Die Datenpunkte verändern sich über einen zufällig gewählten Zeitintervall von drei Monaten, von Bild 1 (Messzeitpunkt 01.08.2007), zu Bild 2 (Messzeitpunkt 24.08.2007), nach Bild 3 (Messzeitpunkt 26.09.2007) und schließlich zu Bild 4 (30.10.2007). Das Ziel dieser Abbildung ist es, dem Leser einen Eindruck zu vermitteln, wie die Animationsfunktion in TimeRider aussieht.

## 5 Usability

*“Usability is the measure of the quality of the user experience when interacting with something - whether a Website, a traditional software application, or any other device the user can operate in some way or another.” - (Jakob Nielsen, 1998)*

In diesem Kapitel wird Usability oder Benutzerfreundlichkeit detailliert beschrieben. Im nachfolgenden Unterkapitel 5.1 wird der Begriff allgemein erläutert und abgegrenzt. Darauf folgt unter 5.2 die Definition des für diese Arbeit relevanten Usability-Begriffs, mit einer kurzen Darstellung der Usability ISO Norm im Unterkapitel 5.2.1. Danach wird im Unterkapitel 5.3 „the usability pope“ (Wirtschaftswoche Magazine, Germany zitiert auf der Homepage der „NN/g“, 2014) Jakob Nielsen vorgestellt, sowie Teile seiner Arbeit zu Benutzerfreundlichkeit, welche eine Grundlage dieser Magisterarbeit bilden. Die Unterkapitel 5.4, 5.5 und 5.6 beschreiben die Usability Kategorien, nach denen TimeRider überprüft wurde.

### 5.1 Allgemein

Das Wort „*Usability*“ kommt vom Begriff „*user friendly*“ (benutzerfreundlich) und wurde Anfang der 80er Jahre geprägt. Unter Usability versteht man die Benutzbarkeit eines Produktes für den Nutzer (Homepage der Universität Rostock, 2002). Usability ist somit eine hypothetische Eigenschaft, die in erster Linie Softwareprodukten zugeschrieben wird, wenn sie z.B. "benutzerfreundlich", "angenehm zu bedienen" und "geeignet zur Erfüllung einer bestimmten Aufgabe" sind (Homepage der Universität Regensburg, 2015).

In dieser Magisterarbeit liegt der Focus auf Usability in der Informatik. Usability ist ein technisches Hauptforschungsgebiet im Bereich der Mensch-Computer Interaktion (*HCI = Human-Computer Interactions*). In der HCI werden verschiedenste Aspekte der Benutzung und Verwendung von Computern durch den Menschen, hauptsächlich im Kontext von interaktiven Informatiksystemen untersucht. Durch die Betrachtungsweise „*Mensch-Computer*“ ist Usability in der Informatik ein interdisziplinäres angewandtes Forschungsgebiet. Aus verschiedenen ausgewählten Forschungsgebieten werden die jeweiligen Ansät-

ze für eine Lösungsfindung zusammengeführt. Des Weiteren ist Usability von den dynamischen Wechselwirkungen der vier prinzipiellen Komponenten (*user, task, tool, environment*) eines „User-Systems“ und nach Situation abhängig, siehe Abbildung 15 (Shackel & Richardson, 1991).

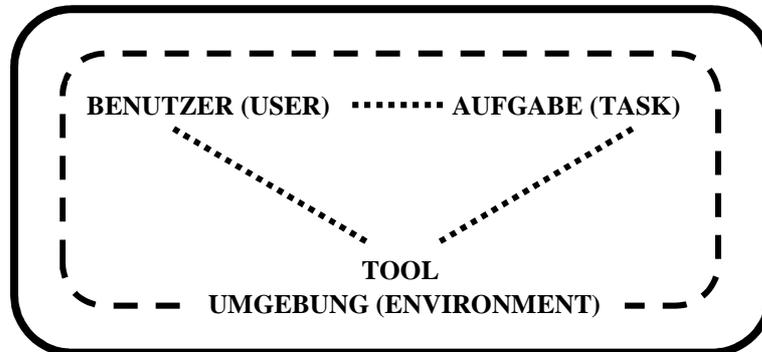


Abbildung 15: Schematische Darstellung der Usability-Wechselwirkung (in Anlehnung an Shackel & Richardson, 1991, p.24)

## 5.2 Definition

*“Usability is the capability in human functional terms to be used easily and effectively by the specified range of users, given specified training and user support, to fulfil the specified range of tasks, within the specified range of environmental scenarios.” - (Shackel & Richardson, 1991).*

Im deutschsprachigen Raum wird Usability mit den Begriffen „Benutzerfreundlichkeit“, „Benutzbarkeit“ und „Gebrauchstauglichkeit“ übersetzt (Homepage der Universität Rostock, 2002). Der Begriff „Benutzerfreundlichkeit“ entstand, als für Computerverkäufer der Mensch als BenutzerIn in den Fokus rückte. Dieser deutsche Begriff ist eigentlich irreführend, da BenutzerInnen keine Maschinen brauchen um freundlich zu diesen zu sein, sondern nur um eine Aktion aus- oder durchzuführen. Bei einem Arbeitsschritt beispielsweise wollen BenutzerInnen, durch eine Maschine diesen durchführen, dabei sollte die Maschine die UserInnen jedoch unterstützen und nicht die Arbeit erschweren, oder sogar im schlimmsten Fall daran hindern. Ebenfalls könnte „Benutzerfreundlichkeit“ den Trugschluss zulassen, dass die Ansprüche der BenutzerInnen mit einem eindimensionalen System, welches mehr oder weniger „freundlich“ ist, beschrieben werden könnten. In der Realität haben verschiedene UserInnen, verschiedene Vorstellungen und Ansprüche, somit kann ein System von eine(r) BenutzerIn als „freundlich“ wahrgenommen werden, während andere UserInnen es als langweilig empfinden (Nielsen, 1994c).

Damit keine Missverständnisse bei dem Forschungsfeld der „Benutzerfreundlichkeit“ entstehen können, werden verschiedene Bezeichnungen verwendet, anbei ein Auszug:

- CHI (*computer-human interaction*),
- HCI (*human-computer interaction* – wird verwendet wenn der Mensch symbolisch an erster Stelle stehen soll),
- UCD (*user-centered design*),
- MMI (*man-machine interface*),
- HMI (*human-machine interface*),
- OMI (*operator-machine interface*),
- UID (*user interface design*),
- HF (*human factors*), Ergonomy (*HF* und *ergonomy* haben ein breiteres Spektrum als HCI. Viele Usability-Methoden sind auf Designs von verschiedenen komplexen und nicht einfach eingängigen Systemen anwendbar. Die Methoden sind aber auch bei eingängigen Systemen anwendbar, welche dadurch noch eingängiger und einfacher zum Erlernen sind.) (Nielsen, 1994c).

Programme oder -systeme werden als benutzerfreundlich gesehen, wenn die Bedienung einfach und stabil gegen Fehlverhalten der UserInnen, sowie ohne Vorkenntnisse leicht erlernbar ist. Benutzerfreundlichkeit ist somit eine subjektive Einschätzung des Nutzers, wobei „Benutzbarkeit“ und „Gebrauchstauglichkeit“ objektive Kriterien sind und sich auf die Funktionalität und Leistungsfähigkeit des Programms beziehen (Homepage der Universität Rostock, 2002).

Aus dem Blickwinkel der angewandten Forschung, ist Usability die Einbindung wissenschaftlicher Erkenntnisse in den Design- und Entwicklungsprozess technischer Gerätschaften (vorwiegend Software). Das Ziel dabei ist es, dem Endanwender ein intuitiv bedienbares Produkt anzubieten (Homepage der Universität Regensburg, 2015).

Usability sollte über die bloße Gestaltung der Benutzerschnittstelle des Produktes hinausgehen und der Entwurf eines Produktes an bestimmte Benutzergruppen angepasst werden. Darüber hinaus sollten die BenutzerInnen in deren Arbeitsabläufen und Arbeitsaufgaben unterstützt werden, um ein effektives, effizientes und zufriedenes Arbeiten zu ermöglichen (Homepage der Universität Rostock, 2002).

Wie bereits erwähnt handelt es sich bei Usability um ein interdisziplinäres Fachgebiet. Damit ein Programm benutzerfreundlich wird, sollten also auch unterschiedlichste Faktoren in Betracht gezogen werden, wie z.B. intuitive Verständlichkeit (einfache Erlernbarkeit), der Aufgabe entsprechend, Erwartungskonformität (Vorwissen oder Verhaltensstere-

otypen berücksichtigen), sowie Fehlerrobustheit (siehe nächstes Kapitel 5.2.1) (Homepage der Universität Rostock, 2002).

### 5.2.1 ISO 9241.110

Die DIN EN ISO 9241 beschreibt die Qualitätsrichtlinien zur Sicherstellung der Ergonomie interaktiver Systeme. Die *International Standards Organisation (ISO)* wählte zunächst den Titel „Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten“, bis dieser 2006 in „Ergonomie der Mensch-System-Interaktion“ geändert wurde, um die Einschränkung auf Büroarbeiten aufzuheben (Homepage des „Handbuch Usability“, 2007).

Der Unterabschnitt 9241.10 wurde 2006 überarbeitet und aktualisiert. Die Bezeichnung wurde auf 9241.110 geändert, mit dem Titel „Grundsätze der Dialoggestaltung“. Der Begriff Benutzerschnittstelle (*User Interface*) ist darin als „*alle Bestandteile eines interaktiven Systems, die Informationen und Steuerelemente zur Verfügung stellen, die für die BenutzerInnen notwendig sind, um eine bestimmte Arbeitsaufgabe mit dem interaktiven System zu erledigen*“ definiert. Ebenfalls werden in 9241.110 die nachfolgenden „*sieben Grundsätze der Dialoggestaltung*“ vorgestellt (Homepage des „Handbuch Usability“, 2007), welche im wesentlichen Usability beschreiben:

- i. Aufgabenangemessenheit („*Ein Dialog ist aufgabenangemessen, wenn er die BenutzerInnen unterstützt, deren Arbeitsaufgabe effektiv und effizient zu erledigen.*“)
- ii. Selbstbeschreibungsfähigkeit („*Ein Dialog ist selbstbeschreibungsfähig, wenn jeder einzelne Dialogschritt durch Rückmeldung des Dialogsystems unmittelbar verständlich ist oder den BenutzerInnen auf Anfrage erklärt wird.*“)
- iii. Erwartungskonformität („*Ein Dialog ist erwartungskonform, wenn er konsistent ist und den Merkmalen der BenutzerInnen entspricht, z.B. den Kenntnissen aus dem Arbeitsgebiet, der Ausbildung und der Erfahrung der BenutzerInnen, sowie den allgemein anerkannten Konventionen.*“)
- iv. Fehlertoleranz („*Ein Dialog ist fehlertolerant, wenn das beabsichtigte Arbeitsergebnis trotz erkennbar fehlerhafter Eingaben, entweder mit keinem oder mit minimalem Korrekturaufwand, durch die BenutzerInnen erreicht werden kann.*“)
- v. Steuerbarkeit („*Ein Dialog ist steuerbar, wenn die BenutzerInnen in der Lage sind, den Dialogablauf zu starten sowie seine Richtung und Geschwindigkeit zu beeinflussen, bis das Ziel erreicht ist.*“)

- vi. Individualisierbarkeit („*Ein Dialog ist individualisierbar, wenn das Dialogsystem Anpassungen an die Erfordernisse der Arbeitsaufgabe, individuelle Vorlieben der BenutzerInnen und Benutzerfähigkeiten zulässt.*“)
- vii. Lernförderlichkeit („*Ein Dialog ist lernförderlich, wenn er die BenutzerInnen beim Erlernen des Dialogsystems unterstützt und anleitet.*“)

### 5.3 Jakob Nielsen – Der Usability Papst<sup>5</sup>

Wenn Usability in einer wissenschaftlichen Arbeit beschrieben wird, muss Herr Jakob Nielsen, Ph.D genannt werden. Er ist gemeinsam mit Dr. Donald A. Norman (ehemaliger Vizepräsident für Forschung bei Apple Computer) Mitbegründer der Nielsen Norman Group (NNg). Nielsen gilt als Gründer der Bewegung „*discount usability engineering*“, welche für schnelle und billige Entwicklungen und Verbesserungen von Benutzerfreundlichkeit (z.B.: Schnittstellen, *Interfaces*) steht. Ebenfalls hat er diverse Usability Methoden entwickelt, unter anderem die „*heuristic evaluation*“ (Beurteilung von Elementen anhand von Usability Kategorien), die „*heuristic estimation*“ (Einschätzung von Verhalten der UserInnen) und „*consistency inspection*“ (Vergleich von Designs und *Interfaces*). Jakob Nielsen besitzt 79 U.S. Patente, die meisten beschreiben einfachere Benutzung des Internets (Homepage der „NN/g“, 2014).

Nielsen verfasste unzählige Bücher und Forschungsberichte zu Usability, die Hauptwerke bilden dabei seine drei Bücher „*Usability Engineering* (1993)“; „*Multimedia and Hypertext: The Internet and Beyond* (1995)“, sowie „*Designing Web Usability: The Practice of Simplicity* (1999)“.

Im Buch „*Usability Engineering*“ von Nielsen (1993), werden Usability Techniken im Detail beschrieben, um verbesserte Produktqualität zu ermöglichen und Designfehler frühzeitig zu erkennen. Nielsen gibt darin auch eine schrittweise Anleitung, welche Methoden in den verschiedenen Phasen eines Entwicklungszyklus zu verwenden sind. Weitere detaillierte Informationen darin sind, wie ein Usability-Test abzuhalten ist und welche besonderen Herausforderungen bei internationaler Benutzerfreundlichkeit zu beachten sind. Auch kosteneffektive Methoden welche vom Entwickler sofort implementiert werden können und Strategien zur Vermeidung der häufigsten Gründe von Verzögerung bei Softwareprojekten sind in dem Buch zu finden (Homepage der „NN/g“, 2014).

Die 2te Auflage des Buchs „*Multimedia and Hypertext: The Internet and Beyond*“ von Nielsen (1995) benutzt Hypertext als weiteren Schritt zu Usability im Internet. Basierend

---

<sup>5</sup> Jakob Nielsen – Der Usability Papst, siehe Homepage der „NN/g“, 2014.

auf einer inhaltlichen Fundierung der Hypertext-Theorie, werden eine Reihe von Strategien darin vorgestellt, wie z.B. Strategien zur Lösung von Informationsüberladung, Erklärungen zu Urheberrecht für UserInnen und EntwicklerInnen, Usability Herausforderungen hinsichtlich Hypertext und Kurz- sowie Langzeit Prognosen hinsichtlich der Zukunft dieses Forschungsgebiets (Homepage der „NN/g“, 2014).

Das Buch „*Designing Web Usability: The Practice of Simplicity*“ von Nielsen (1999) ist ein Handbuch für Benutzerfreundlichkeit im Internet und basiert auf den Ergebnissen von sechs Jahren Forschung und ungefähr 400 Usability Test mit UserInnen aus verschiedensten Fachrichtungen. Von dem Inhalt einer Webseite über das Webseitendesign für einfachere Navigation bis zu BenutzerInnen mit Einschränkungen, beschreibt dieses Buch Wegweiser und Anleitungen, wie jed(e) beliebige WebnutzerIn in jeder Situation angesprochen und zum Verweilen animiert werden kann (Homepage der „NN/g“, 2014).

Usability bezieht sich auf alle Aspekte eines Systems, welches mit Menschen in irgendeine Art und Weise interagieren kann und bezieht sich auch auf die Installation und Wartung eines Systems. Ein Computer ohne „*User Interface*“ Komponente ist kaum zu finden, selbst bei einer technischen Anlage, in der zwischen zwei Produktionscomputern Daten übertragen werden, wird normalerweise eine Schnittstellenoberfläche inkludiert, um Fehler in der Verbindung anzuzeigen (Mulligan, et al., 1991 zitiert in Nielsen, 1994c).

### 5.3.1 Definition von Usability<sup>6</sup>

Usability ist nicht ein einfacher, eindimensionaler Bestandteil einer Benutzerschnittstelle. Benutzerfreundlichkeit besteht aus mehreren Komponenten, welche traditionell die folgenden fünf sind:

- *learnability* (Lernfähigkeit): Das System sollte einfach zu lernen sein, damit keine langen Einarbeitungszeiten notwendig sind und die BenutzerInnen das System schnell zum Arbeiten verwenden können (Nielsen, 1994c).
- *efficiency* (Effizienz): Das System sollte effizient zu verwenden sein, damit nachdem die BenutzerInnen das System erlernt haben, ein hoher Grad an Produktivität möglich ist (Nielsen, 1994c).
- *memorability* (Denkwürdigkeit): Das System sollte „erinnerungswürdig“ sein. Ein(e) durchschnittliche BenutzerIn sollte, selbst wenn das Programm länger nicht verwendet wurde, das Programm jederzeit und ohne Einschulung oder Einarbeitungsphase wieder benutzen können (Nielsen, 1994c).

---

<sup>6</sup> Definition von Usability nach Jakob Nielsen (Nielsen, 1994c).

- *errors* (Fehler): Das System sollte eine niedrige Fehlerrate haben, damit während der Verwendung des Systems so wenige Fehler wie möglich durch die BenutzerInnen verursacht werden können. Wenn es zu Fehlern kommen sollte, muss es den UserInnen möglich sein diese wieder einfach zu beheben. Schwere Fehler welche unabsehbare Folgen verursachen können, dürfen gar nicht erst entstehen können (Nielsen, 1994c).
- *satisfaction* (Zufriedenheit): Das System sollte angenehm zu verwenden sein, damit während UserInnen das System verwenden, diese subjektiv zufriedengestellt werden (Nielsen, 1994c).

Usability wird typischerweise durch mehrere Tests der UserInnen gemessen, wobei die Tests auf die endgültigen Ziel-BenutzerInnen abgestimmt werden sollten. Bei den Tests der UserInnen, führen die ProbandInnen einige spezifische Aufgaben mit dem zu untersuchenden System durch. Gemessen wird Usability relativ zu bestimmten UserInnen und zu bestimmten Aufgaben. Daher kann es vorkommen, dass für dasselbe untersuchte System bei verschiedenen Aufgaben mit verschiedenen UserInnen, die Usability-Charakteristiken sich unterscheiden. Daher beginnt jede Usability-Messung mit der Definition einer repräsentativen Aufstellung von Testaufgaben und wie diese in Relation zu den verschiedenen Usability Attributen gemessen werden können (Nielsen, 1994c).

### 5.3.2 Die zehn Usability Heuristiken<sup>7</sup>

In diesem Unterkapitel werden kurz die zehn Prinzipien für Interaktionsdesign für die „*heuristic estimation*“ oder „Heuristik-Einschätzung“ (Einschätzung des Verhaltens von UserInnen) von Jakob Nielsen beschrieben. Die zehn Prinzipien werden „*heuristics*“ genannt, da Sie großzügige angelegte Regeln sind und keine spezifischen Usability Richtlinien (Nielsen, 1994b).

Die „Heuristik-Einschätzung“ wurde von Rolf Molich und Jakob Nielsen 1990 in Ihren Arbeiten Molich & Nielsen, 1990 und Nielsen & Molich, 1990 beschrieben (Molich & Nielsen, 1990 und Nielsen & Molich, 1990 zitiert auf der Homepage der „NN/g“, 1995). Danach verfeinerte Nielsen seine Heuristik basierend auf 249 Usability-Problemen (Nielsen, 1994a), um die unten stehende Liste von Heuristiken mit maximaler Aussagekraft zu erstellen (Nielsen, 1994b).

#### 1. *Visibility of system status* – Sichtbarkeit des System Status

Das System sollte laufend durch angemessenes Feedback innerhalb eines angebrachten Zeitrahmens, die BenutzerInnen darüber informieren was vor sich geht.

---

<sup>7</sup> Die zehn Usability Heuristiken des Jakob Nielsen (Nielsen, 1994b).

## **2. *Match between system and the real world* – Das Spiel zwischen System und der realen Welt**

Das System sollte die Sprache der BenutzerInnen sprechen, mit Wörtern, Phrasen und Konzepten, welche den UserInnen vertraut sind, im Gegensatz zur System-orientierten Nomenklatur. Reale Konventionen helfen, Informationen in einer natürlichen und logischen Reihenfolge darzustellen.

## **3. *User control and freedom* – Benutzerkontrolle und Freiheit**

BenutzerInnen könnten unbeabsichtigt Systemfunktionen verwenden, daher muss es einen „Notausstieg“ geben, damit der ungewünschte Zustand beendet werden kann ohne extra langen Anweisungen folgen zu müssen. Empfohlen wird auch die Implementierung der Funktionen „undo/redo“ also „Rückgängig/Nochmal“.

## **4. *Consistency and standards* – Konsistenz und Standards**

BenutzerInnen sollten sich niemals wundern müssen. Verschiedene Worte, Situationen oder Aktionen, sollten immer dasselbe meinen. Plattform Konventionen müssen eingehalten werden.

## **5. *Error prevention* - Fehlervorsorge**

Besser als eine gute und hilfreiche Fehlermeldung, ist ein sorgfältiges Design welches verhindert, dass Probleme in erster Linie überhaupt auftreten. Entweder werden zu Fehlerneigende Bedingungen vermieden oder es wird nach diesen laufend überprüft und die UserInnen bekommen eine zu bestätigende Meldung bevor die Aktion ausgeführt werden kann.

## **6. *Recognition rather than recall* – Wiedererkennung anstatt erinnern**

Minimiere die Anzahl der Dinge, die sich die BenutzerInnen merken müssen, durch die Bereitstellung von sichtbaren Objekten, Aktionen und Optionen. Die UserInnen sollten sich nicht an Informationen von einem Teil der Interaktion mit dem System in einem anderen Teil erinnern müssen. Anleitungen zur Verwendung des Systems sollten sichtbar oder einfach, wann immer notwendig abrufbar sein.

## **7. *Flexibility and efficiency of use* – Flexibilität und Effektivität der Nutzung**

Abkürzungen, welche den durchschnittlichen UserInnen nicht bekannt sind, ermöglichen den erfahrenen UserInnen oft eine schnellere Interaktion mit dem System. Das System sollte diese Abkürzungen daher sowohl unerfahrenen und erfahreneren BenutzerInnen zur

Verfügung stellen. BenutzerInnen sollten außerdem in der Lage sein, häufig verwendete Aktionen anpassen zu können.

### **8. *Aesthetic and minimalist design* – Ästhetisches und minimalistisches Design**

Dialoge sollten niemals Informationen enthalten, welche irrelevant oder selten verwendbar sind. Jede zusätzliche Einheit von Information in einem Dialog lenkt von den relevanten Informationen ab und verringert deren relative Sichtbarkeit.

### **9. *Help users recognize, diagnose, and recover from errors* – Hilf den UserInnen Fehler zu erkennen, zu analysieren und zu beheben.**

Fehlermeldungen sollten in einfacher und verständlicher Sprache (im Klartext und keine Codes) den Fehler exakt beschreiben und eine konstruktive Lösung anbieten.

### **10. *Help and documentation* – Hilfe und Dokumentation**

Ein System sollte prinzipiell auch ohne Dokumentation benutzbar sein, es könnte aber trotzdem notwendig sein Hilfestellungen und Dokumentierung anzubieten. Jegliche dieser Informationen sollten einfach zu durchsuchen sein, auf die Aufgaben der BenutzerInnen fokussiert sein, eine Liste mit konkret auszuführenden Schritten enthalten und nicht zu groß sein.

## **5.4 Die acht goldenen Regeln des Interface Design**

Im Buch Shneiderman und Plaisant (2009) werden im Kapitel 2.3.4 „acht goldene Regeln“ für Interface Designs beschrieben, welche auf die meisten interaktiven Systeme anwendbar sind. Diese Richtlinien wurden aus über 30 Jahren Erfahrung, Auswertung und Designs von speziellen Domains gewonnen und verfeinert. Die erste Version dieser Liste wurde 1985 vorgestellt und seit dem auch von Jakob Nielsen und Jeff Johnson, sowie anderen Usability-Spezialisten weiterentwickelt und ergänzt (Shneiderman & Plaisant, 2009).

Die folgenden acht goldenen Regeln sollten für jedes Projekt, jede Softwareplattform und jede Software neu interpretiert, weiterentwickelt und erweitert werden. Die Regeln haben Ihre Grenzen, bieten jedoch eine Grundlage für Mobilfunkgeräte-, PC-Desktop- und Webdesigner (Shneiderman & Plaisant, 2009).

#### **(i) Streben nach Einheitlichkeit (*strive for consistency*):**

Ähnliche Situationen sollten einheitliche Abläufe von Aktionen erfordern, einheitliche Terminologie sollte in Eingabeaufforderungen verwendet werden, sowie in Menüs und

Hilfsmenüs. Ebenfalls sollten einheitliche Farben, Aufbau, Großschreibungen, Schriftarten, usw. sollten durchgehend verwendet werden. Ausnahmen sollten nachvollziehbar und limitiert sein, wie z.B. eine zusätzliche Bestätigung des „Löschen-Kommandos“ oder der zweifachen Passwordeingabe (Passwortbestätigung/ -wiederholung) (Shneiderman & Plaisant, 2009).

(ii) Streben nach universeller Usability (*cater to universal usability*):

Verstehen der Bedürfnisse von unterschiedlichen BenutzerInnen und Designs für Formbarkeit, Unterstützung der Transformation von Inhalten. Bei Berücksichtigung der Unterschiede zwischen AnfängerInnen und ExpertInnen, der Altersgruppen, von körperlichen und oder geistigen Einschränkungen und der technologischen Vielfalt, bereichern das Spektrum des Designs. Sowohl die Realisierung von Funktionen für Unerfahrene, z.B. durch Erklärungen, und Funktionen für Experten, wie Shortcuts und schnellere Steuerungsoptionen, können das Interface Design bereichern und die wahrgenommene Systemqualität verbessern (Shneiderman & Plaisant, 2009).

(iii) Informatives Feedback anbieten (*offer informative feedback*):

Jede Aktion der BenutzerInnen sollte eine Reaktion oder Rückmeldung hervorrufen. Für häufige und untergeordnete Aktionen kann die Antwort dezent sein, wobei bei bedeutenden und seltenen Aktionen sollten deutliche Rückmeldungen vom Programm kommen. Veränderungen sollten eindeutig durch visuelle Präsentation der Objekte welche von Interesse sind, dargestellt werden (Shneiderman & Plaisant, 2009).

(iv) Abgeschlossene Dialoge designen (*design dialogs to yield closure*):

Sequenzen von Aktionen sollten in Gruppen organisiert werden, mit einen Anfang, einen Mittelteil und einen Ende. Informationen und Feedback bei einem Abschluss mehrerer zusammenhängender Aktionen erzeugen ein Gefühl der Befriedigung, der Vollendung, eine Gefühl der Erleichterung, ein Zeichen weitere geplante Aktionen nicht durchzuführen, sowie um auf weitere mögliche Aktionen hinzuweisen. Ein Beispiel dafür sind die Online-Shops, welche einen Besucher von der Auswahl eines Produktes bis zum Bezahlvorgang leiten und am Ende mittels einer eindeutigen Bestätigungsseite abschließen und somit auf die Vollständigkeit der Transaktion hinweisen (Shneiderman & Plaisant, 2009).

(v) Fehlervermeidung (*prevent errors*):

Ein Ziel des Designs eines Systems sollte sein, den UserInnen, so gut wie es nur geht, nicht zu ermöglichen schwerwiegende Fehler zu verursachen. Beispiele sind ausgegraute Menüelemente damit die UserInnen sie nicht verwenden können, wenn nicht geplant und die gesperrte Eingabe von Buchstaben in numerischen Feldern. Sollten die UserInnen dennoch

einen Error verursachen, sollte das Interface diesen automatisch entdecken und eine simple, konstruktive und spezifische Anweisungen zur Wiederherstellung anbieten. Als Beispiel hierfür sollten BenutzerInnen in einer Eingabemaske bei einem Fehler eines Feldes nicht die gesamte Maske neu ausfüllen müssen, sondern eben nur die fehlerhafte Eingabe korrigieren. Versehentliche Aktionen sollten den Systemstatus unberührt verbleiben lassen oder das Interface sollte Anweisungen geben wie man den Ursprungszustand wiederherstellen kann (Shneiderman & Plaisant, 2009).

- (vi) Einfaches rückgängig machen von Aktionen ermöglichen (*permit easy reversal of actions*):

Soweit als möglich sollten Aktionen auch rückgängig gemacht werden können. Diese Funktion lindert Ängste, weil die BenutzerInnen wissen, dass Fehler rückgängig gemacht werden können und ermutigt ungewohnte Optionen zu erkunden. Die Rückgängigmachung sollte für verschieden Einheiten vorgesehen werden, wie einfache Aktionen, Dateneingabefunktionen oder komplette Gruppen von Aktionen, wie die Eingabe in eine Datenmaske vom Namen bis zur Adresse (Shneiderman & Plaisant, 2009).

- (vii) Interne Kontrollüberzeugung unterstützen (*support internal locus of control*):

Erfahrene BenutzerInnen verlangen vor allem das Gefühl, dass Sie das Interface steuern und dass dieses auf Ihre Aktionen reagiert. Diese BenutzerInnen wollen keine Überraschungen oder Veränderungen von gewohnten Verhalten und sie sind schnell verärgert über langweilige Dateneingabesequenzen, umständliche Vorgänge um Informationen zu erlangen und das gesetzte Ziel aufgrund des Designs nicht zu erreichen (Shneiderman & Plaisant, 2009).

- (viii) Belastungen des Kurzzeitgedächtnisses reduzieren (*reduce short-term memory load*):

Der Beschränkung der menschlichen Kurzzeitgedächtnisspanne für Informationsverarbeitung (Millersche Zahl (Miller, 1956): Ein Mensch behält gleichzeitig nur  $7 \pm 2$  Informationseinheiten (*chunks*) im Kurzzeitgedächtnis) sollte ins Design einfließen, um Interfaces zu vermeiden in denen BenutzerInnen sich die Informationen von einer Bildschirmdarstellung merken müssen, aber in einer anderen erst verwenden können (z.B. Webseitenadressen sollten immer sichtbar bleiben) (Shneiderman & Plaisant, 2009).

## 5.5 Das Informationsvisualisierungs-Mantra<sup>8</sup>

In diesem Kapitel werden die 10 Heuristiken aus dem Mantra für Informationsvisualisierung von Forsell & Johansson (2010) vorgestellt. Diese 10 Kategorien werden in dieser Magisterarbeit bei der Kategorisierung von Usability Problemen des zu untersuchenden Programms verwendet um die Fehleranalyse zu vereinfachen, siehe Kapitel 8.5.4. Die Basis des Mantras bildet die Arbeit von Nielsen (Nielsen, 1994a zitiert in Forsell & Johansson, 2010). Im Detail Niensens Überlegung eine endliche und definitive Menge an Heuristiken zu definieren, welche die entdeckten Probleme des zu untersuchenden Programms größtmöglich erklärend umfasst. Die wichtigsten Usability Heuristiken für Informationsvisualisierung wurden dabei als Vorlage verwendet, wie Informationskodierung, Abbildung der Daten (*information coding, mapping*), minimale Aktionen der UserInnen (*minimal actions*), „Abschließbarkeit“ (*exibility*, Anzahl der möglichen Lösungswege für eine Aufgabe), Orientierung und Hilfe (*orientation and help*), räumliche Anordnung (*spatial organization*, z.B. Verteilung der Elemente auf dem Bildschirm), Konsistenz (*consistency*), Wiedererkennung anstatt erinnern (*recognition rather than recall*), (BenutzerInnen-) Führung (*prompting*, alle Mittel welche den UserInnen helfen alternative Wege zu finden um Aufgaben zu erledigen), Entfernung alles Fremden (*remove the extraneous*, Ablenkung durch unnötige Information), Datensatzreduzierung (*data set reduction*) (Forsell & Johansson, 2010).

Die meisten heuristischen Sets vor Forsell & Johansson (2010) umfassten die je nach angewandter Situation wichtigen Aspekte, aber keine der Heuristiken konnte als allgemeine Regel für jegliche Informationsvisualisierungen herangezogen werden. Des Weiteren gibt es noch keinen Konsens in der Wissenschaft darüber, welche Art von Heuristik hilfreich ist für die Einschätzung von spezifischen Aspekten bei komplexen interaktiven visuellen Darstellungen in Informationsvisualisierungen (Zuk, et al., 2006 zitiert in Forsell & Johansson, 2010).

Damit eine repräsentative Stichprobe möglich war, welche einen großen Bereich von generellen Problemen abdeckt, untersuchten Forsell & Johansson (2010) ein breites Spektrum an Visualisierungstools. Dazu gehörten von einfachen Konzeptvisualisierungen (Forsell, et al., 2006 zitiert in Forsell & Johansson, 2010) über vollfunktionale Prototypen (Lundblad, 2008 sowie Vrotsou, 2009 zitiert in Forsell & Johansson, 2010) bis zu Produkten welche bereits in wirtschaftlicher Nutzung waren (*Spotfire and map of the market* zitiert in Forsell & Johansson, 2010). (Forsell & Johansson, 2010).

Für die Erstellung eines allgemein gültigen Mantras für Informationsvisualisierung verwendeten Forsell & Johansson daher eine Datenbank von Problemen, welche die beiden

---

<sup>8</sup> Das Informationsvisualisierungs-Mantra nach Forsell & Johansson (2010).

Forscher aus 74 Usability-Problemen aus früheren Usability-Studien gesammelt haben (Forsell, et al., 2006 und Lundblad, 2008 sowie Vrotsou, 2009 zitiert in Forsell & Johansson, 2010). Diese 74 unikalen Probleme bilden eine repräsentative Stichprobe, da ein großer Bereich von generellen Problemen abgedeckt wird, welche in Informationsvisualisierungs-Systemen gefunden werden. Generell entstehen meist Probleme bei der virtuellen Repräsentation (z.B. präzises dekodieren von Daten, lesen von Informationselementen, usw.), der Interaktionsmechanismen (z.B. Daten filtern mittels Interaktion, Daten Interaktionsmechanismen wie Zoomen, etc.) und der Einfachheit und Support bei Erforschung von Daten (z.B. finden von Erkundungsabläufen, Datenvergleiche, usw.). Die Probleme und Ihre Ausprägungen sind im Detail nicht von Interesse. Wichtig ist nur, dass bei der Benutzung eines Systems dasselbe Problem auftritt und den UserInnen auch auffällt, denn damit ist klar ersichtlich, welche Fehler bei einem „Redesign“ des Systems unbedingt und wichtig zu entfernen sind (Forsell & Johansson, 2010).

In Forsell & Johansson (2010) wurden 6 ProbandInnen, welche alle ExpertInnen auf dem Gebiet der Informationsvisualisierung und oder HCI an der Linköping University waren, beauftragt, 74 Usability-Probleme anhand von 6 Heuristik-Sets zu bewerten. Diese sechs Heuristik-Gruppen (A bis F) wurden aus anderen Publikationen ausgewählt, welche zusammengerechnet insgesamt 63 Heuristiken umfassten. Für eine detaillierte Beschreibung der Heuristiken sei an dieser Stelle auf die original Publikation der Gruppen verwiesen (siehe Beschreibung der Gruppen weiter unten), da die Details den Rahmen dieser Magisterarbeit sprengen würden. Diese sechs Gruppen von Heuristiken sind Bestandteil verschiedenster wissenschaftlicher Arbeiten. Diese Arbeiten wurden von den Autoren gezielt ausgewählt, da sie eine große Vielfalt von verschiedensten Perspektiven hinsichtlich Usability mit Bezug auf Informationsvisualisierung enthielten, alle bereits publiziert waren und um die Anzahl der Sets zu begrenzen, damit die ProbandInnen die Studie hinsichtlich Arbeitspensum bewältigen können (Forsell & Johansson, 2010). Diese Sets sind wie folgt:

**A:** *Visual Information Seeking Mantra* (Craft & Cairns, 2005 zitiert in Forsell & Johansson, 2010): 7 Heuristiken.

**B:** *Ergonomic Criteria for Hierarchical Information Visualization Techniques* (Freitas, et al., 2002 zitiert in Forsell & Johansson, 2010): 9 Heuristiken.

**C:** *Ten Usability Heuristics* (Nielsen, 1994b und Homepage der „NN/g“, 1995 zitiert in Forsell & Johansson, 2010): 10 Heuristiken.

**D:** *Perceptual and Cognitive Heuristics* (Zuk & Carpendale 2006 zitiert in Forsell & Johansson, 2010): 13 Heuristiken.

**E:** *Ergonomic Criteria Interactive Systems* (Scapin & Bastien, 1997 zitiert in Forsell & Johansson, 2010): 18 Heuristiken.

**F:** *Knowledge Tasks* (Amar & Stasko 2004 zitiert in Forsell & Johansson, 2010): 6 Heuristiken.

Die Gruppenbezeichnung **A** bis **F** wird in weiterer Folge auch in der restlichen Magisterarbeit verwendet.

Basierend auf der Bewertung der sechs Heuristik Sets durch die ProbandInnen, wurden die 10 Heuristiken mit dem breitesten Erklärungsspektrum von Forsell & Johansson (2010) ausgewählt (*top ten*). Nachstehend sind diese 10 Heuristiken beschrieben, wobei der Großbuchstabe beschreibt in welchem Set (A-F) die Heuristik vorkommt und die Zahl (1 bis 18) beschreibt, welche der Heuristiken innerhalb eines jeweiligen Sets genau gewählt wurde (Forsell & Johansson, 2010):

**B5 = Information coding (Informationskodierung):** Diese Heuristik basiert darauf, dass die Wahrnehmung von Informationen direkt von der Abbildung (fachsprachlich „Mapping“, Homepage des „Duden“, 2015c) von Datenelementen auf visuelle Objekte abhängig ist. Daher sollten realistische Charakteristiken, Techniken oder zusätzliche Symbole verwendet werden, damit die Wahrnehmung einer gewünschten Information verbessert wird (Forsell & Johansson, 2010).

**E7 = Minimal actions (Minimale Aktionen):** Diese Heuristik beschreibt, dass um ein Ziel zu erreichen oder eine Aufgabe zu erfüllen das Arbeitspensum auf ein mögliches Minimum von Aktionen zu reduzieren ist (Forsell & Johansson, 2010).

**E11 = Flexibility (Vielseitigkeit):** Diese Heuristik verweist darauf mehrere Lösungswege anzubieten, damit UserInnen-spezifische Arbeitsstrategien, Gewohnheiten und Voraussetzungen zum Erreichen eines vorgegebenes Ziels möglich sind (Forsell & Johansson, 2010).

**B7 = Orientation and help (Orientierung und Hilfe):** Diese Heuristik sieht vor Funktionen zu inkludieren, wie Hilfe für die Einstellung des Detailgrads, Rückgängigmachen oder Wiederholen von Aktionen (*redo/undo*) und das Anzeigen von zusätzlichen und weiterführenden Informationen (Forsell & Johansson, 2010).

**B3 = Spatial organization (räumliche Organisation):** Diese Heuristik berücksichtigt die Orientierung der BenutzerInnen im Informationsraum, die Verteilung von Elementen im Layout, Präzision und Leserlichkeit, effiziente Raumnutzung und Verzerrungen von visuellen Elementen (Forsell & Johansson, 2010).

**E16 = Consistency (Konsistenz):** Diese Heuristik referenziert darauf ob das Design bei ähnlichen Zusammenhängen auch gleich ist (Beständigkeit des Frontenddesigns) und in verschiedenen Kontexten verschieden ist (keine Verwechslungsgefahr durch Ähnlichkeiten) (Forsell & Johansson, 2010).

**C6 = Recognition rather than recall (Wiedererkennung anstatt erinnern):** Diese Heuristik besagt, dass BenutzerInnen sich nicht an viel Information erinnern sollen, um eine Aufgabe durchzuführen (Forsell & Johansson, 2010).

**E1 = Prompting (Aufforderungen):** Diese Heuristik umfasst die Hilfestellungen, welche den UserInnen alle alternativen möglichen Aktionen abhängig vom Kontext anzeigen (Forsell & Johansson, 2010).

**D10 = Remove the extraneous (Entfernung aller Elemente, welche mit der Sache nichts zu tun haben):** Diese Heuristik bezieht sich auf zusätzliche Informationen welche für die UserInnen eine Ablenkung sein könnten und die Aufmerksamkeit fehlleiten, indem die eigentlichen Daten nicht gesehen oder Vergleiche nicht erkennbar sind (Forsell & Johansson, 2010).

**B9 = Data set reduction (Datensatzreduktion):** Diese Heuristik beschreibt Funktionen zur Reduktion von Datensätzen, deren Effizienz und einfache Handhabung (Forsell & Johansson, 2010).

Diese 10 ausgewählten und zusammengefassten Heuristiken (B3, B5, B7, B9, C6, D1, E1, E7, E11, E16) basieren auf den jeweiligen original Beschreibungen der jeweiligen Autoren. Bei Interesse an einer umfassenderen Beschreibung von diesen Heuristiken, sei an dieser Stelle aus Platzgründen nochmals auf die originalen Publikationen verwiesen. Die vorher angeführte Liste enthält vier aus neun Heuristiken des Sets „**B: Ergonomic Criteria for Hierarchical Information Visualization Techniques** (Freitas, et al., 2002 zitiert in Forsell & Johansson, 2010)“, eine aus 10 Heuristiken des Sets „**C: Ten Usability Heuristics** (Nielsen, 1994b und Homepage der „NN/g“, 1995 zitiert in Forsell & Johansson, 2010)“, ebenfalls eine aus 13 Heuristiken des Sets „**D: Perceptual and Cognitive Heuristics** (Zuk & Carpendale 2006 zitiert in Forsell & Johansson, 2010)“ und vier aus 18 Heuristiken aus dem Set „**E: Ergonomic Criteria Interactive Systems** (Scapin & Bastien, 1997 zitiert in Forsell & Johansson, 2010)“ (Forsell & Johansson, 2010).

Abschließend ist noch zu erwähnen, dass diese 10 Heuristiken auf zu dieser Zeit bekannten Usability Problemen basieren und für die Auswertung der Untersuchung dieser Magisterarbeit angewandt wurde, siehe Kapitel 8.5.4. Die Autoren empfehlen Ihre 10 Heuristiken bei zukünftigen Heuristik Evaluierungen zu verwenden, um bei der Suche und Auswertung

von neuen Usability Problemen deren Anwendbarkeit zu verifizieren (Forsell & Johansson, 2010).

## 5.6 Die sieben Interaktionskategorien<sup>9</sup>

*“The importance of interaction and the need for its further study seem undisputed.” - (Yi, et al., 2007).*

In diesem Kapitel werden die auf den Bedürfnissen der UserInnen basierenden sieben Kategorien von Interaktionstechniken nach Yi, et al. (2007) vorgestellt, da sie in dieser Masterarbeit zur Anwendung kommen (siehe Kapitel 8.5.5). Die Kategorisierungen sollen einen Einblick in den Interaktionsprozess während der Benutzung von Informationsvisualisierungssystemen geben und für die Fehleranalyse transparenter machen (Yi, et al., 2007).

Yi, et al. (2007) sehen im Fokus von Informationsvisualisierungssystemen zwei Hauptkomponenten: Repräsentation und Interaktion. Die repräsentative Komponente hat ihren Ursprung im Forschungsfeld Computergraphik und umfasst das Mapping von Daten zur Repräsentation und wie diese am Display gerendert werden. Die interaktive Komponente hat hingegen ihren Ursprung im Forschungsfeld der HCI und beschäftigt sich mit dem Dialog zwischen den UserInnen und dem System, während die BenutzerInnen den Datensatz erforschen um neue Erkenntnisse zu gewinnen. Obwohl sich diese zwei Komponenten ergänzen, lag der Fokus bei wissenschaftlichen Arbeiten auf der Repräsentations-Seite. Die interaktive Komponente wird nur in einigen wenigen Papers beschrieben (vgl. Buja, et al., 1991 und Dix & Ellis, 1998 und Kosara, et al., 2003 und Tweedie, 1997 zitiert in Yi, et al., 2007). Dennoch ist Interaktion ein essentieller Teil von Informationsvisualisierung, da ohne Interaktionen eine Informationsvisualisierungstechnik zu einem statischen Bild oder zu autonom animierten Bildern verkommt (vgl. *InfoCanvas* in Miller & Stasko, 2001 zitiert in Yi, et al., 2007). Statische Bilder haben hingegen den Vorteil, dass sie einen analytischen und ausdrucksstarken Wert haben (vgl. Bertin, 1983 und Norman, 1993 und Tufte, 1990 zitiert in Yi, et al., 2007). Diese Stärke schwindet jedoch proportional mit steigender Größe des Datensets und deren Variablen. Interaktion findet jedoch auch bei statischen Bildern statt, wenn z.B. bei einem Diplomarbeitposter der Betrachter einige Interaktionen durchführt, wie das Poster rotieren, näher herangehen oder weiter entfernen und dabei ansehen, Notizen auf das Poster schreiben oder kleben (Haftnotiz). Spence (2007, zitiert in Yi, et al., 2007) beschreibt in seiner Arbeit den Begriff „passive Interaktion“, gemeint damit ist das von der BenutzerInnen mentale Model der Daten, welches durch sie ebenfalls bereits gedanklich verändert oder weiterentwickelt werden kann (Spence, 2007 zitiert in Yi, et al., 2007).

---

<sup>9</sup> Die sieben Interaktionskategorien nach Yi, et al. (2007).

Nun folgenden die sieben Kategorien nach Yi, et al. (2007) im Detail, sowie Beispiele dieser Kategorien anhand von bereits implementierten Programmfunktionen des in dieser Masterarbeit zu untersuchenden Programms TimeRider. Die sieben Kategorien basieren auf der Absicht, dass ein(e) UserIn vorsätzlich mit einem Informationsvisualisierungssystem zu interagiert und der vorher beschriebenen Interaktionskomponente und deren Prozess bei der Benutzung des Systems, zwischen den UserInnen und den repräsentierten Daten (Yi, et al., 2007):

### **1. *SELECT* - mark something as interesting (markiere etwas als interessant):**

Diese Interaktionskategorie umfasst jede Aktion der BenutzerInnen, die eine Möglichkeit bietet, Datenpunkte die im Fokus des Interesses liegen zu markieren und dadurch einfach verfolgbar zumachen. Wenn in ein und derselben Anzeige mehrere Datenpunkte vorhanden sind oder wenn Darstellungen verändert werden, wird es schwierig für UserInnen den Datenpunkten von Interesse zu folgen. Indem Datenpunkte visuell unverwechselbar gemacht werden können, können die BenutzerInnen die Übersicht selbst in großen Datensätzen und oder wechselnden Darstellungen behalten (Yi, et al., 2007).

Ein Beispiel für die *Select* Technik ist, wenn die UserInnen einen Datenpunkt markieren können und dieser Punkt durch die Anwahl dann anders eingefärbt bleibt (z.B. Datenpunkt wird durchs anklicken automatisch eindeutig eingefärbt, während die anderen Datenpunkte ihre ursprüngliche Farbe beibehalten). Es können Datenpunkte anhand nach bestimmten Kriterien markiert werden, welche sich dann farblich verändern und einfach zu verfolgen sind. Auch wenn die Anordnung der Datenpunkte verändert wird, können die BenutzerInnen danach ganz einfach mit einem Blick die Position des markierten Datenpunktes feststellen. Diese Funktion wurde jedoch nicht in TimeRider implementiert.

### **2. *EXPLORE* - show me something else (zeige mir etwas anderes):**

Diese Interaktionskategorie beschreibt die Betrachtung von Unterkategorien von Datensätzen durch die BenutzerInnen. Wenn UserInnen in einem Informationsvisualisierungssystem Daten ansehen, sehen sie meist nur einen Teil oder Ausschnitt auf einmal. Dies kann an mehreren Gründen oder an einer Kombination aus diesen liegen. Zum Beispiel an einem zu großen Maßstab des Datensatzes, durch die Limitierung der Sicht und oder des Bildschirms, durch fundamentale kognitive und Wahrnehmungsgrenzen der menschlichen Informationsverarbeitung. BenutzerInnen eines Informationsvisualisierungssystems sehen sich typischerweise eine Unterkategorie des Datensatzes an, um die Daten zu Verstehen und neue Erkenntnisse zu erlangen und dann erst sehen sie sich weitere Daten an. Bei *Explore*-Interaktionen werden nicht notwendigerweise die Daten komplett verändert die auch gerade betrachtet werden, häufiger kommt es vor, dass neue Daten in die gerade bestehende Ansicht kommen während andere entfernt werden (Yi, et al., 2007).

Ein Beispiel für eine *Explore* Technik ist das sogenannte „*panning*“. Nach Yi, et al. (2007) ist mit „*panning*“ entweder die Bewegung einer Kamera über eine dargestellte Szene oder die Bewegung einer Szene während die Kamera still steht gemeint. Bei Softwareanwendungen ist „*panning*“ eine speziell implementierte Möglichkeit, bei der die UserInnen entweder die Szene greifen können (linke Maustaste, bei Standardeinstellung der Rechtshänder-Option, drücken und durchgängig gedrückt halten) und mit der Bewegung der Computermaus diese bewegen oder einfach mittels Mauszeiger über eine bedienbare Bildlaufleiste (oder auch *scrollbar*) den Ablauf verändern können. Siehe hierfür Abbildung 16 in der „*panning*“ in TimeRider mittels Zeitschieber umgesetzt wurde und die Funktion der Zeitachse wie eine Scrollbar eines Media Players auf einem PC verhält.

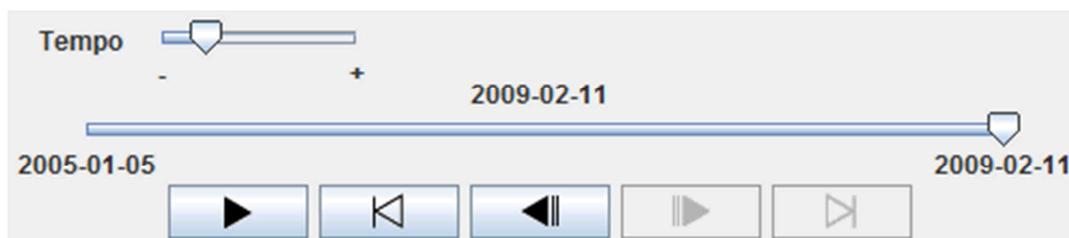


Abbildung 16: Explore Technik in TimeRider

### 3. *RECONFIGURE* - *show me a different arrangement* (zeig mir ein anderes Arrangement):

Diese Interaktionskategorie beinhaltet alle Änderungen welche die räumliche Anordnung der Darstellung betreffen, durch verschiedene Perspektiven auf den Datensatz. Einer der essentiellen Anwendungszwecke der Informationsvisualisierung ist es versteckte Charakteristiken von Daten aufzudecken und die Beziehung zwischen ihnen. Eine „gute“ Darstellung berücksichtigt dies zwar, eine einzelne Repräsentation bietet jedoch meist selten eine hinreichend Perspektiven. Wenn *Reconfigure*-Interaktionstechniken angeboten werden, dann können die UserInnen die Anordnung der Datenelemente oder deren Ausrichtung ändern, um verschiedene Perspektiven auf den Datensatz zu ermöglichen (Yi, et al., 2007).

Als Beispiel für eine *Reconfigure* Technik, ist die Änderung der Attribute welche der X- und Y-Achsen zugewiesen sind. Bei einer Veränderung werden auch die zu untersuchenden Sets der Merkmale und der Variablen des gesamten Datensatzes verändert. Somit können sich auch Zusammenhänge zwischen Datenpunkten ändern und dadurch verschiedene Perspektiven ermöglichen (Yi, et al., 2007). Siehe dazu Abbildung 17, wo diese Änderung der zugrundeliegenden Werte der X- und Y-Achse, mittels „dropdown“ Menüs bereits realisiert wurde.

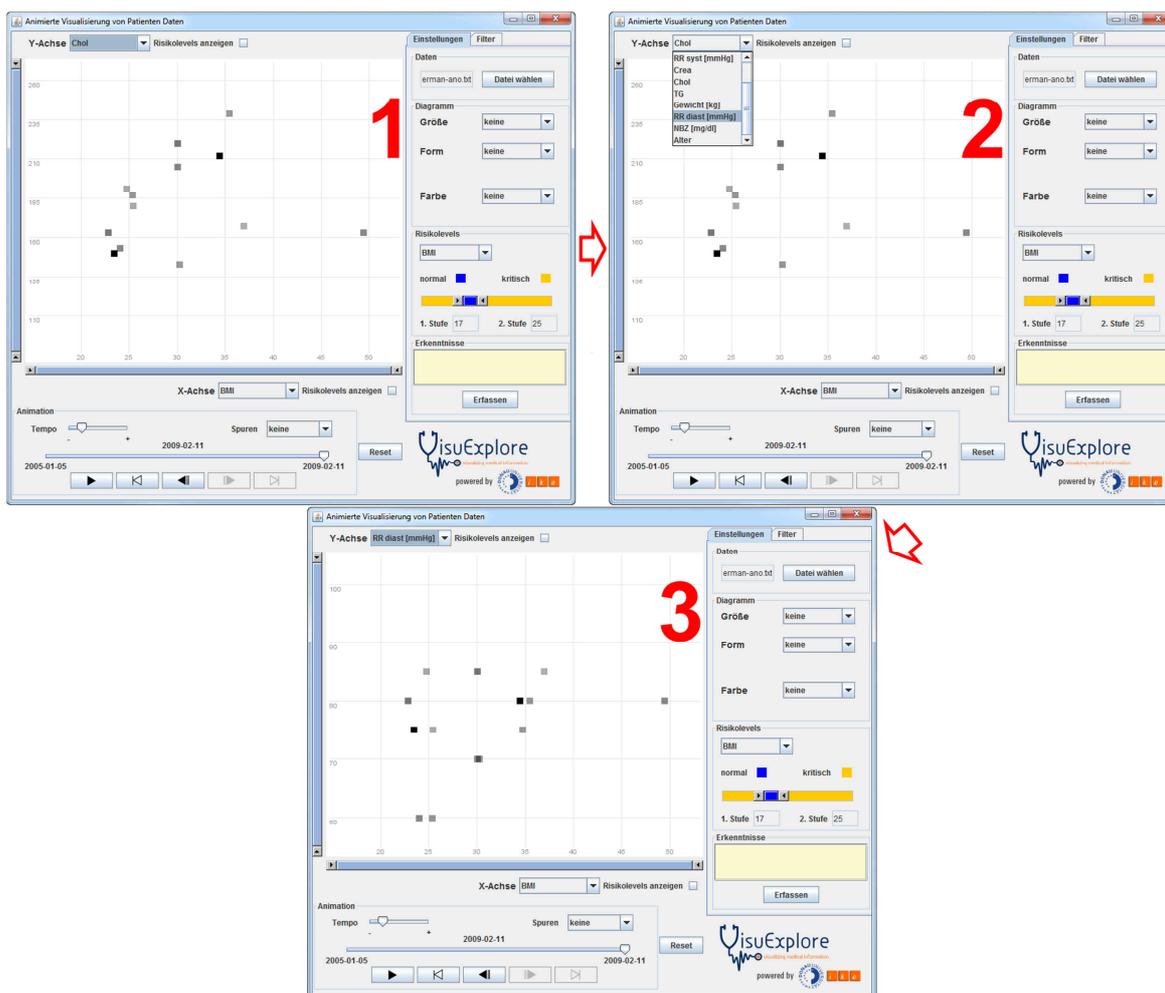


Abbildung 17: Reconfigure Technik in TimeRider<sup>10</sup>

#### 4. ENCODE - show me a different representation (zeig mir eine andere Darstellung):

Diese Interaktionskategorie umfasst alle Techniken die ein(e) UserIn hat, um fundamentale visuelle Darstellungen der Daten inklusive visuellen Erscheinungsbild (z.B. Farbe, Größe und Form) eines jeden Datenelements zu verändern. Visuelle Elemente in Informationsvisualisierungssystemen spielen eine wichtige Rolle, da sie sowohl die präattentive Wahrnehmung (Vor-Aufmerksamkeits-Wahrnehmung, siehe detaillierten Exkurs im Anhang C) beeinflussen können, als auch direkt das Verständnis der UserInnen für Beziehungen und Verteilung der Daten (Yi, et al., 2007).

Ein Beispiel für die *Encode* Technik, ist die Verwendung von Farbschattierungen bei Höheninformationen, damit die BenutzerInnen die Höhe eines Objekts (z.B. die Höhe eines Berges) besser einschätzen können ohne die räumliche Anordnung der Karte zu verändern.

<sup>10</sup> *Reconfigure* Technik in TimeRider: Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die Möglichkeit die Parameter in den Achsen „dropdown“ Menüs zu verändern. In Bild 1 sieht man die Ausgangssituation. In Bild 2 sieht man wie eine andere Achsen-Variable ausgewählt wird. In Bild 3 dann die daraus resultierende neue Darstellung der Datenpunkte auf den Achsen.

Ebenfalls unter *Encoding* Techniken fallen die Veränderbarkeit der Größe der Datenpunkte, deren Orientierung und deren Schrift und Form (Yi, et al., 2007). In TimeRider wurden drei Techniken implementiert: „Größe, Form, Farbe“. Durch diese einfachen Veränderungen wie die Daten dargestellt werden, z.B. durch die Veränderung der Farbcodierung des Datensatzes (siehe dafür Abbildung 18), erwarten sich die BenutzerInnen neue Zusammenhänge der Daten zu entdecken.

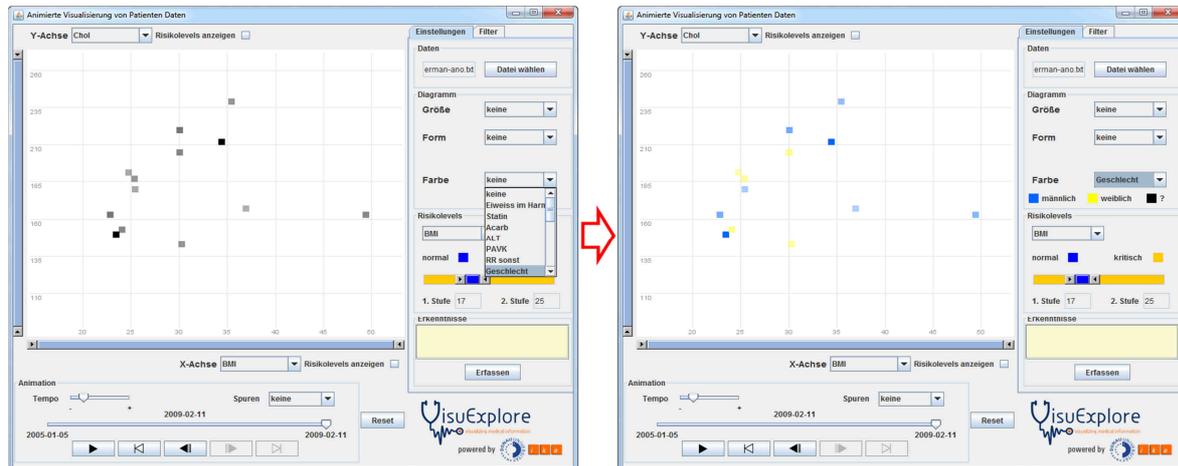


Abbildung 18: Encode Technik in TimeRider<sup>11</sup>

### 5. ABSTRACT/ELABORATE - show me more or less detail (zeig mir mehr oder weniger Details):

Diese Interaktionskategorie beinhaltet jede Aktion der BenutzerInnen, die eine Möglichkeit bietet den Grad der Abstraktion der Datendarstellung anzupassen. Diese Interaktionsarten erlauben es BenutzerInnen die Darstellung von der Übersicht bis ins Detail zu ändern, damit einzelne Falldaten und alle Ebenen dazwischen eingesehen werden können. Die Absicht der UserInnen liegt jeweils darin, entweder einen breiten kontextbasierten Überblick der zu untersuchenden Daten zu erlangen, oder die individuelle Werte eines speziellen Falls zu untersuchen (Yi, et al., 2007).

Ein Beispiel für eine *Abstract/Elaborate* Technik sind einfache Tool-Tipp Interaktionstechniken, sogenanntes „*mouseover*“ oder „*hovering*“. Diese zeigen Detailinformationen an, wenn der Mauszeiger über Datenpunkte bewegt wird (siehe Abbildung 19). Das klassische Beispiel für diese Interaktionskategorie ist jedoch das „*Zoomen*“. Damit sind sowohl das geometrische wie das semantische Zoomen gemeint (siehe Abbildung 20). Durch das Zoomen können BenutzerInnen einfach den Maßstab der Darstellung ändern indem sie, entweder um einen Überblick über den gesamten Datensatz und dessen Kontext zu erlangen herauszoomen, oder um die Details des Datensatzes in den Fokus zu bringen hinein-

<sup>11</sup> *Encode* Technik in TimeRider: Die rote Kennzeichnung wurde vom Autor dieser Arbeit nachträglich eingefügt und zeigt die Änderung der Farbcodierung des Datensatzes in TimeRider. Im linken Bild ist die Auswahl der Variablen im dropdown Menü „Farbe“ zu sehen. Das rechte Bild zeigt dann, die sichtbar veränderte Farbcodierung der Datenpunkte.

zoomen. Die Intention der UserInnen beim Hineinzoomen ist es, die bestehende Ansicht einer Informationssammlung/ -ansicht zu ändern, um eine vergrößerte oder detailliertere Ansicht einer bestimmten Region anzuzeigen (Gotz & Zhou, 2008). Der Datensatz sollte also beim Zoomen sich selbst jedoch nicht ändern, sondern nur ein Teil der gewählten Darstellung (Yi, et al., 2007).

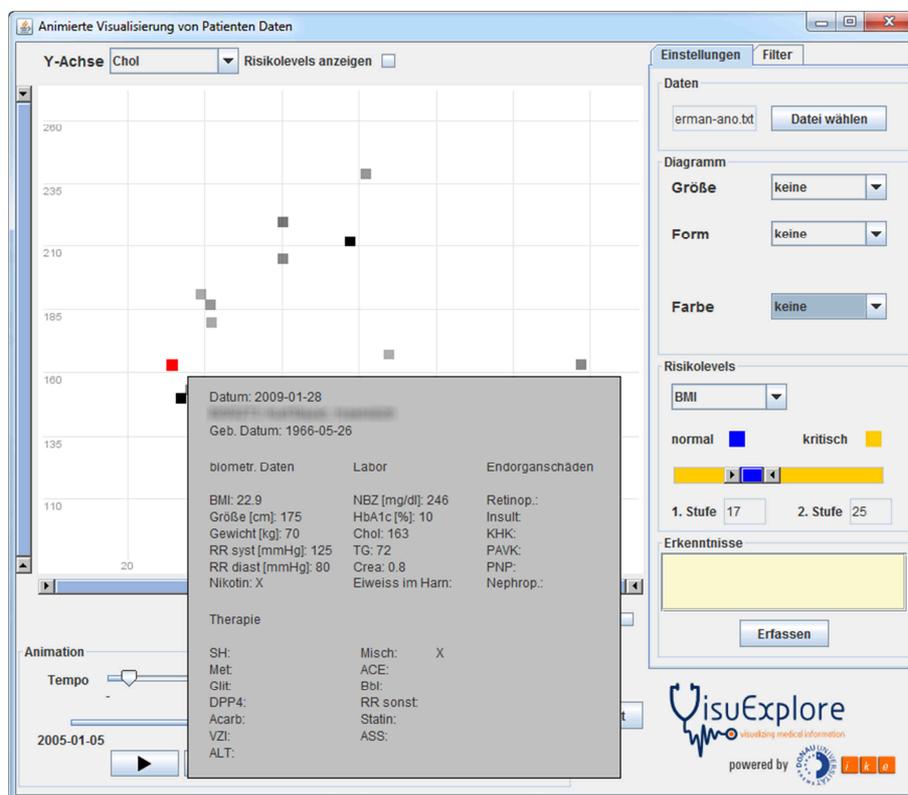


Abbildung 19: Abstracte/Elaborate Technik - „mouseover“ in TimeRider

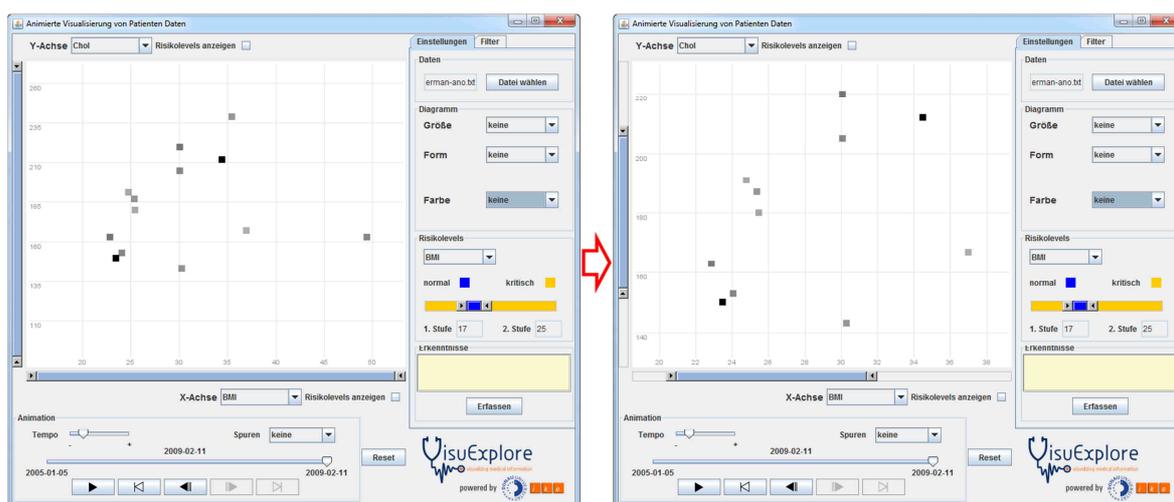


Abbildung 20: Abstracte/Elaborate-Technik - „zoomen“ in TimeRider<sup>12</sup>

<sup>12</sup> Abstracte/Elaborate Technik in TimeRider: Die rote Kennzeichnung wurde vom Autor dieser Arbeit nachträglich eingefügt und zeigt im linken Bild die Ausgangssituation und im rechten Bild den Datensatz nach dem „Hineinzoomen“, z.B. mittels Mausrad. Anhand der Schieberegler der Achsen sieht die betrachtende Person, dass bereits und wie weit gezoomt wurde.

## 6. *FILTER: show me something conditionally* (zeig mir etwas unter gewissen Bedingungen)

Diese Interaktionskategorie umfasst alle Interaktionstechniken, welche den UserInnen anhand von spezifischen Bedingungen ermöglicht die Präsentation der Datenpunkte zu verändern. In dieser Kategorie spezifizieren die BenutzerInnen eine gewünschte Anzahl von Bedingungen, damit nur jene Daten angezeigt werden, welche unter die Kriterien der festgelegten Bedingungen fallen. Datenpunkte welche von der aktivierten Bedingung nicht erfasst werden können oder solche die die Bedingung nicht erfüllen, werden entweder nicht oder verändert angezeigt. Die aktuellen Daten verbleiben üblicherweise unverändert, damit bei der Rückgängigmachung (Reset) der gewählten Kriterien, die versteckten/ausgeblendeten oder verändert dargestellten Datenpunkte wiederhergestellt werden können. Die UserInnen ändern nicht die Perspektive auf die Daten, sondern verändern nur die Darstellungskonditionen der Daten (Yi, et al., 2007).

Ein einfaches Beispiel für eine *Filter* Technik wird in Abbildung 21 gezeigt. Dies ist ein sogenanntes „dropdown“ Menü, also ein Kontextmenü eines Programms welches durch draufklicken mit dem Mauszeiger aktiviert wird und dann unmittelbar „ausklappt“, indem die ProbandInnen verschiedenste Werte oder Variablen auswählen können.

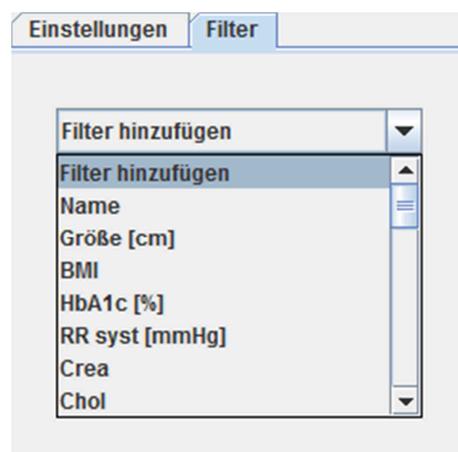


Abbildung 21: Filter Technik in TimeRider<sup>13</sup>

## 7. *CONNECT: show me related items* (zeig mir verwandte Datenpunkte):

Diese Interaktionskategorie beinhaltet jede Aktion, welche einerseits zum Herausheben von Verbindungen und Beziehung zwischen den angezeigten Datenpunkten und andererseits zum Anzeigen von versteckten aber punktuell spezifisch relevanten Datenpunkten, verwendet wird (Yi, et al., 2007).

<sup>13</sup> Filter Technik in TimeRider :Aufgeklapptes dropdown Menü des Karteireiters Filter im Programm TimeRider.

Ein Beispiel für eine *Connect* Technik ist das Hervorheben von verbundenen oder benachbarten Datenpunkten durch das Bewegen des Mauszeigers über einen Datenpunkt. Dabei soll jedoch nicht die gesamte Darstellung verändert werden, sondern nur die Daten die von Interesse sind. Dadurch können auch Daten oder Zusammenhänge enthüllt werden, welche zuvor nicht sichtbar waren (Yi, et al., 2007). In TimeRider wurde die *Connect* Technik, einerseits durch die sogenannten „Traces“ (Spuren) umgesetzt. In Bild 1 der Abbildung 22 sieht man die aktivierte Spuren-Darstellung „nur letzte Untersuchung“, wobei hier nur der letzte Untersuchungszeitpunkt als großes ausgefülltes und schwarz gefärbtes Quadrat dargestellt wurde und die vorhergehenden Untersuchungen mittels einer durchgehenden Linie visualisiert sind. Im rechten Bild 2 ist die Spuren-Darstellung „alle vergangenen Untersuchungen“ aktiviert. Diese Darstellung bewirkt, dass der letzte Messzeitpunkt, das große schwarze Quadrat, mit dem die vorhergehenden Messzeitpunkte, die kleinen schwarzen Quadrate, mittels einer durchgehenden Linie verbunden wurden.

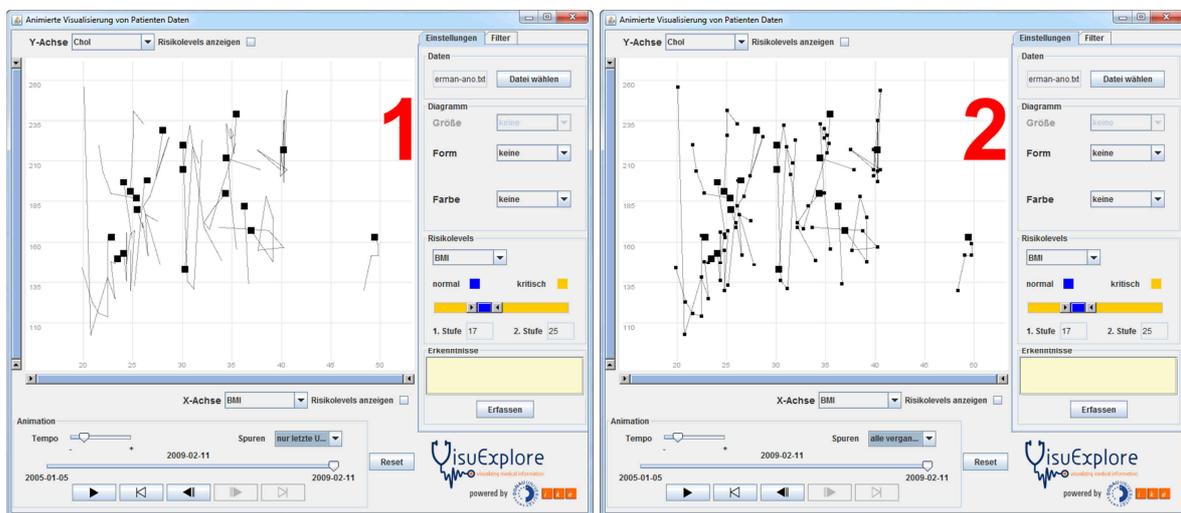


Abbildung 22: Connect Technik - „traces“ in TimeRider

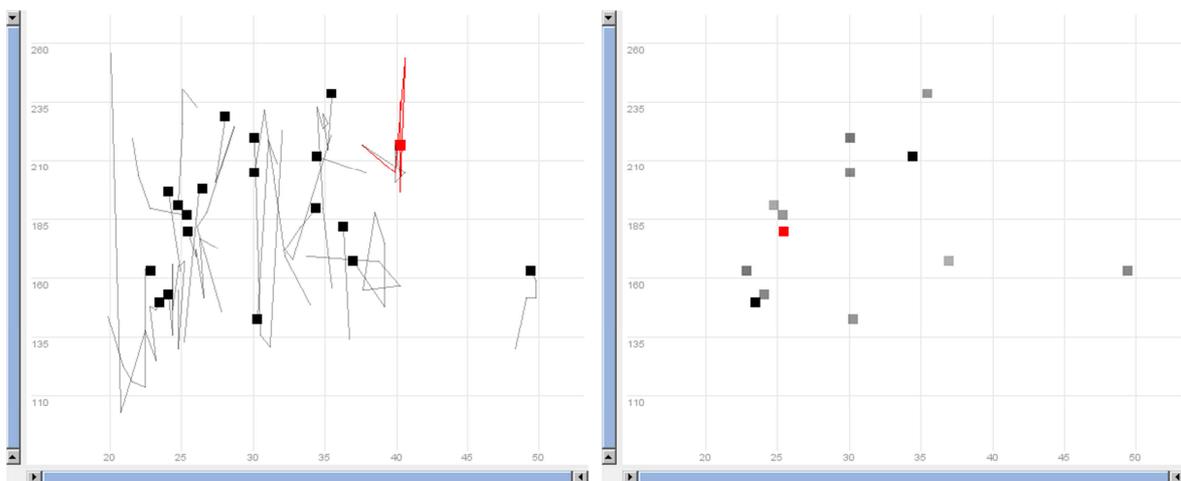


Abbildung 23: Connect Technik - „highlighting“ in TimeRider

Andererseits wurde die *Connect* Technik in TimeRider auch als Hervorhebung („highlighting“) von Datenpunkten implementiert, sowohl während die Traces aktiviert sind, als

auch im normalen Scatter Plot (in TimeRider wird der Trace oder der Datenpunkt, bei einem „mouseover“ rot eingefärbt, siehe Abbildung 23).

### 5.6.1 Weitere Interaktionstechniken

Andere Interaktionstechniken in Informationsvisualisierungssystemen auf die hier nicht näher eingegangen wurde, da diese oder so ähnliche Operationen nicht in TimeRider implementiert wurden, ist eine breite Palette an Techniken, welche in den meistern interaktiven Applikationen vorkommen, wie z.B. (Yi, et al., 2007):

- „undo/redo“ Funktionen (Rückgängig/Nochmal), welche es den UserInnen erlauben vorwärts und rückwärts zu vorherigen Systemstatus zugehen (z.B., *undo*, *redo*, *history* und *reset*) (Yi, et al., 2007).
- Veränderungs-Konfigurationen: Das sind Techniken die es den UserInnen erlauben verschiedenste Einstellungen und Konfigurationen am System vorzunehmen (z.B. bei einem Browser in der Menüleiste im „dropdown“-Menü „Extras“ und dann unter „Einstellungen“)

Da solche Operationen heute in den verschiedensten Arten von Applikationen standardmäßig verwendet werden und somit auch keine besonderen Auswirkungen auf Informationsvisualisierungssysteme zu erwarten sind, wurden sie nicht berücksichtigt. Dies bedeutet aber keine Abwertung dieser Techniken, im Gegenteil sie haben eine feste Größe als nützlicher interaktive Möglichkeiten innerhalb der Informationsvisualisierung (Yi, et al., 2007).

Laut Pohl, et al. (2012) reflektieren die sieben Kategorien nach Yi, et. al. (2007) die Aktivitäten der BenutzerInnen und nicht die des Systems, daher eignen sich diese für, wie in dieser Arbeit durchgeführten Untersuchungen. Diese Kategorien wurden erstellt, damit verschiedene Informationsvisualisierungssysteme verglichen werden können (Pohl, et al., 2012).

Zum Abschluss wird in Yi, et al. (2007) nochmals auf die Untrennbarkeit von Darstellung und Interaktion hingewiesen, da beide den UserInnen helfen Informationen besser zu verstehen und gleichzeitig die kognitiven Belastungen senken, frei nach dem Motto: „Ein Bild sagt mehr als tausend Worte“ (Fred R. Barnard - Homepage der “The history of a picture's worth”, 2015) (Yi, et al., 2007).

## 6 Thinking Aloud

In diesem Kapitel 6 wird die Methode des „*Thinking Aloud*“ beschrieben und im Unterkapitel 6.1 die Untersuchungsmethode dieser Magisterarbeit, das „*Thinking Aloud Protokoll*“ nach Boren & Ramey (2000) im Detail selbst vorgestellt.

Die Methode der Wahl in dieser Magisterarbeit, zur Beantwortung der definierten Forschungsfragen ist „*Thinking Aloud*“ und wird in diesem Kapitel beschrieben. Diese Methode entwickelte sich aus dem Behaviorismus heraus. Laut dem Duden Online (2015d) bedeutet „Behaviorismus“: „*Richtung der amerikanischen Verhaltensforschung, die nur direkt beobachtbares Geschehen als Gegenstand wissenschaftlicher Psychologie zulässt*“.

Zu Beginn des 20. Jahrhunderts lag der Fokus der amerikanischen Psychologen in der Verhaltensforschung von menschlichen Individuen. Die inneren geistigen Abläufe wurden als „nicht-beobachtbar“ bezeichnet, da sie laut Watson (1913) für die Psychologie irrelevant sind. Der Begriff „Behaviorismus“ wurde erstmals in eben dieser Arbeit „*Psychology as the behaviorist views it*“ von Watson (1913) erwähnt. Watson attackiert darin die damals in der Psychologie gebräuchliche Methode der Untersuchung des Bewusstseins. Er kritisierte die introspektive Methode samt ihrer Resultate und argumentierte, dass die Psychologie als eine Naturwissenschaft ohne Introspektion und ohne mentale Konstrukte auskommt (Ericsson & Simon, 1980). Nach Watson (1913) werden objektive Beobachtungen erhoben und kategorisiert, ohne Rücksicht auf die Struktur oder Funktionsweise des Geistes, indem jedes Verhalten entweder als Reiz (*input*) oder Reaktion (*output*) gesehen wird. Watson bezeichnete als Reiz, jede Veränderung in der umgebenden Außenwelt oder innerhalb des menschlichen Individuums. Die Voraussetzung war, dass der Reiz auf physiologischen Vorgängen beruht, z.B. bei „Wassermangel im menschlichen Körper“ empfindet der Mensch „Durst“ und die menschliche Reaktion darauf ist „Wasser trinken“. Im Behaviorismus wird der menschliche Geist als Black-Box<sup>14</sup> gesehen, welche die Zwischenschritte der Wahrnehmung (innere Prozesse im menschlichen Gehirn) nicht abbildet und eine ä-

---

<sup>14</sup> Laut der Homepage des *Oxford English Dictionary Online* (2015) ist eine „*black box*“: [...] Ein Gerät welches komplexe Funktionen ausführt, aber dessen interne Mechanismen möglicherweise nicht leicht be-sichtigt oder verstanden werden können, (daher) jede Komponente eines Systems welche in Bezug auf die Wechselbeziehung Eingabe und Ausgabe spezifiziert ist (siehe Abbildung 24).

berliche Beobachtung der innerseelischen Vorgänge nicht zulässt. Ein Reiz (*input*) führt in die Black-Box und der menschliche Organismus reagiert darauf (*output*) (siehe Abbildung 24). Eine behavioristische Untersuchung analysiert daher nur die beobachtbaren Zusammenhänge zwischen Reiz (*stimuli*) und Reaktion (*response*), daher wird dieser Teilbereich der Psychologie auch als Reiz-Reaktions-Psychologie bzw. *stimulus-response-psychology* (*S-R psychology*) bezeichnet (Homepage der Universität Duisburg-Essen, 2015).



Abbildung 24: behavioristisches Wahrnehmungsbild<sup>15</sup>

Ein wichtiger Punkt ist, dass Watson (1913 zitiert in Ericsson & Simon, 1980) seine Kritik nicht gegen die Interpretation von verbal berichteter Information, sondern gegen die analytischen Methoden der klassischen Introspektion gerichtet hatte. Watson hat somit einen ersten nachweislichen Grundstein von Thinking Aloud gelegt, da es nur von Interesse ist welcher Reiz welches Resultat erzeugt. Für Thinking Aloud würde dies bedeuten, nur die Aussagen der ProbandInnen sind interessant und welcher Umstand zu dieser Aussage geführt hat. Auf Computersoftware bezogen wäre dann zu sagen, es ist nur wichtig welche Aktion der UserInnen oder jene mit dem Programm zu einer Aussage geführt haben.

Zwanzig Jahre nach Watson (1913) beschrieb Woodworth (1938 zitiert in Ericsson & Simon, 1980) „verbale Berichte“ (*verbal reports*) und wies dezidiert auf den Unterschied zwischen „Gedanken beschreiben“ (*describing thoughts*) und die „Gedanken zum Ausdruck bringen“ (*expressing thoughts*) hin (Ericsson & Simon, 1980).

In Ericsson & Simon (1980) wurde das erste Mal über verbale Berichte mit Fokus auf die Thinking Aloud Methode beschrieben. Es wurden jedoch keine Studien mit nicht klar definierten Aufgaben oder medizinische Studien zur Recherche verwendet. Ericsson & Simon (1980) beschrieben und interpretierten alle wichtigen historischen Arbeiten, welche zu „*verbal reports*“ geführt haben. Daraus entwickelten sie ein eigenständiges Thinking Aloud Protokoll, welches strikte Rahmenbedingungen für den gesamten Ablauf vorsieht. Boren & Ramey (2000) schlagen eine freiere Gestaltung des „Thinking Aloud“ Prozesses vor, wobei generell angenommen wird, dass das Gesprochene den Denkprozess beim Lösen einer Aufgabenstellung widerspiegelt. Da dies das Thinking Aloud Protokoll ist, welches in dieser Magisterarbeit verwendet wurde, wird im nächsten Kapitel 6.1 näher darauf eingegangen.

<sup>15</sup> Diese Darstellung zeigt das behavioristische Wahrnehmungsbild, indem der menschliche Organismus als Black-Box dargestellt wird (vgl. Homepage der Universität Duisburg-Essen, 2015).

In Trafton, et al. (2000) wird festgestellt, dass der Redefluss als Reflektion des momentan stattfindenden kognitiven Prozesses angesehen werden kann. Nach einer Auswertung können die Informationen über die Prozesse und die Darstellung zur Verbesserung oder weiteren Arbeit verwendet werden. In Kraemer & Ummelen (2004) gehen die Autoren soweit, die Thinking Aloud Methode als eine populäre und effektive Methode für Usability Tests zu bezeichnen. Das Verfahren erzeugt nützliche Informationen über die BenutzerInnen, welche mit einer bestimmten Applikation interagieren. Die hauptsächlich qualitativen Daten enthalten Hinweise, welche Teile oder Bereiche des Programms den BenutzerInnen Schwierigkeiten bereiten und fließen bei Bedarf in die nächste Aktualisierung des Programms mit ein.

In den Studien von Friel, et al. (2001), Trickett und Trafton (2006) und Ratwani, et al. (2008) werden zusätzlich zu den aufgezeichneten und analysierten Daten eines Thinking Aloud Protokolls, verschiedene Kodierungsschemata vorgestellt. Alle Kodierungsmodelle sind direkt vom Datensatz und dem verwendeten System abhängig und müssen je nach Einsatz immer individuell angepasst werden, damit die daraus resultierenden Daten auf Verhaltensmuster analysieren werden können. Darüber hinaus wurden unterschiedliche Kodierungsschemata, zur Analyse von gesammelten Thinking Aloud Daten vorgestellt. Das Kodierungsmodell selbst ist vom Datensatz und dem verwendeten System abhängig und muss daher individuell angepasst werden. Die daraus resultierenden Daten werden dann auf Muster analysiert. Diese Technik als manueller Ansatz wird von Gotz & Zhou (2008) in Ihrem Paper „*Characterizing Users' Visual Analytic Activity for Insight Provenance*“ verwendet und beschrieben.

In Pohl, et al. (2012) wurden anhand von zwei verschiedenen Informationsvisualisierungssystemen sowohl Thinking Aloud Studien durchgeführt, um Usability Probleme im Programm aufzudecken, als auch Logfiles analysiert, um Verhaltensmuster zu beobachten. Diese Muster sind ein Hinweis auf die Strategie die BenutzerInnen wählen, um eine Aufgabe zu lösen (Pohl, et. al., 2012). Im Gegensatz zu dieser Studie, in welcher zwei Informationsvisualisierungssysteme anhand ihrer Logfiles analysiert und verglichen werden, wird in dieser Magisterarbeit nur eine spezifische Visualisierungsprogramm (TimeRider), anhand verbalisierter Gedanken mittels der Thinking Aloud Methode nach Boren & Ramey (2000) untersucht. Die Logfiles dienen in dieser Arbeit nur als Auflistung der Usability und gegebenenfalls der Utility Fehler. In Pohl, et al. (2012) hatte die Thinking Aloud Studie außerdem das Ziel, anhand der Logfiles generelle Interaktionsmuster der UserInnen bei zwei unterschiedlichen Programmen zu finden und in weiterer Folge, ob daraus Muster in Gedankenprozessen auch abgeleitet werden können. Die Definition der Thinking Aloud Methode besagt, dass die laut ausgesprochenen Gedanken die des Problemlösungsprozesses widerspiegeln. Das in dieser Magisterarbeit verwendete Thinking Aloud Protokoll nach Boren & Ramey (2000) wird im nachfolgenden Kapitel 6.1 beschrieben.

## 6.1 Thinking Aloud Protokoll nach Boren & Ramey (2000)

Thinking Aloud Protokolle werden in der Literatur (Boren & Ramey, 2000) als am meist verwendete Methode für Usability Test genannt. In Boren & Ramey (2000) wird der Unterschied zwischen Theorie und Praxis, des bis dahin am meisten verwendeten Thinking Aloud Protokolls nach Ericsson & Simon (1993) untersucht. Dabei fanden die Autoren heraus, dass die meisten publizierten Usability Analysen von den eigentlichen theoretischen Theoremen von Ericsson & Simon (1993) abweichen, obwohl deren Werk „*Protocol Analysis: Verbal Reports as Data*“ explizit als Grundlage angegeben wurde. In diesem Unterkapitel wird der modifizierte Ansatz von Boren & Ramey (2000) beschrieben und es werden an passenden Stellen, Beispiele aus dem Usability Test von TimeRider dem gegenübergestellt.

Wie auch in dem Test der dieser Magisterarbeit zugrunde liegt, bitten die meisten Usability Tester deren ProbandInnen „laut zu denken“. Laut Nielsen (Homepage der „NN/g“, 2015) ist die Thinking Aloud Methode sogar „[...] die einzige und beste Usability Engineering Methode“. Es gibt jedoch keine detaillierte Beschreibung in der Usability Literatur für theoretische Regeln zur praktischen Anwendung von Thinking Aloud. Ein Grund für die Abweichung von der Theorie nach Ericsson & Simon (1993) könnte an den komplexen und strikten Vorgaben liegen (Boren & Ramey, 2000).

In Boren & Ramey (2000) werden die drei verschiedenen Levels von Ericsson & Simon (1993) definiert, wobei der vom Ersten bis zum Dritten die Relevanz der Daten stark nachlässt. Welche Daten berücksichtigt werden sollten, hängt dabei von den Vorgängen im Kurzzeitgedächtnis (*short term memory* = STM) ab. Nachfolgend stehen die drei Kategorien inklusive kurzer Erklärung wie in Boren & Ramey (2000) beschrieben. Jegliche weitere Daten, die von dieser Kategorisierung nicht erfasst werden können, gelten bei Ericsson & Simon (1993) als unzulässig (bewertende Urteile, Bewusstseinsstrom, ...):

- (i) Level 1: Formulierungen, welche gedanklich nicht erst umgewandelt werden müssen um formuliert zu werden. Dies betrifft z.B. Zahlen bei einer Rechenaufgabe, da Zahlen auf dieselbe Art formuliert werden können, wie sie im Kurzzeitgedächtnis kodiert sind. Level 1 ist die verlässlichste Art von Formulierungen unter Ericsson & Simon (1993), aber oft schwer zu erhalten.
- (ii) Level 2: Formulierungen, welche erst transformiert werden müssen, damit sie während einer Bearbeitung einer Aufgabe formuliert werden können. Beispiele dafür wären Bilder oder abstrakte Konzepte, welche formuliert werden sollen während der Lösung einer Aufgabe, solange keine anderen kognitiven Prozesse zwischen

Kurzzeitgedächtnis und Formulierung benötigt werden. Level 2 wird im Modell von Ericsson & Simon (1993) ebenfalls als zuverlässige Daten akzeptiert.

- (iii) Level 3: Formulierungen welche weitere kognitive Prozesse für die Aufgabenlösung oder Formulierung benötigen. Darunter fallen z.B. die Verwendung von geistigen Filterprozessen, wie „formulieren Sie nur diese Informationen ...“. Dadurch kommt es zu einem Informationsaustausch zwischen Kurzzeitgedächtnis und Langzeitgedächtnis und dies könnte die Daten verfälschen. Ebenfalls fallen darunter jegliche Beeinflussung, wie Kommentare oder Aufforderungen der ModeratorInnen, welche die nachfolgenden Daten ebenfalls verfälschen könnte.

Ebenfalls identifizierten Boren & Ramey (2000) in Ihrer Arbeit vier Thinking Aloud Anweisungen von Ericsson & Simon (1993), welche zur Durchführung von Usability Tests verwendet werden können:

- 1) Es sollten nur „harte“ Daten gesammelt und analysiert werden. Dazu zählen die vorher beschriebenen Daten der Levels eins und zwei.
- 2) Die ProbandInnen müssen vorab detailliert in Thinking Aloud eingeschult werden. Die ModeratorInnen sollten durch so wenig Ablenkung wie möglich, die Testperson ermutigen, konstant „laut zu denken“, durch laute nicht notwendigerweise zusammenhängende Aussprache, als wenn niemand im Raum anwesend wäre. Darüber hinaus sollen die ProbandInnen darauf hingewiesen werden, dass die Erinnerung jedes Mal vorgenommen wird, wenn die Testperson aufhört „laut zu denken“. Des Weiteren sollten TeilnehmerInnen dies vorab üben.
- 3) Es sollten die ProbandInnen daran erinnert werden „laut zu denken“. Nach einer vorher definierten Wartezeit und bei schweigenden ProbandInnen, sollten die ModeratorInnen die ProbandInnen auffordern wieder „laut zu denken“, dies sollte so kurz und unbestimmt als möglich durchgeführt werden. Die ModeratorInnen sollten weder eine persönliche Beziehung, noch Erwartungen in den ProbandInnen wecken. Eine erfolgreiche Erinnerung sollte in einer sofortigen Wiederaufnahme von Thinking Aloud enden. Ericsson & Simon (1993) empfehlen als Erinnerungsaufforderung ein simples: „Reden Sie weiter“ („*Keep going*“, „*Go on*“).
- 4) In keinen Fall eingreifen. Nach Beginn einer Aufgabe sollte wenn überhaupt nur die Erinnerungsaufforderung „Reden Sie weiter“ als Interaktion stattfinden. Jede andere Interaktion könnte ablenken, selbst neutral Fragen oder Kommentare der ModeratorInnen können schon ablenken.

Da diese Regeln der Realität von „*software usability*“ Untersuchungen wenig dienlich ist, begannen Boren & Ramey (2000) ausgehend von den strikten Rahmenbedingungen nach

Ericsson & Simon (1993) eine eigene Methode zu entwickeln. Sie beschrieben als erste die gefundenen Unterschiede zwischen der Theorie und den von Ihnen untersuchten Usability Tests. Daraus entwickelten Boren & Ramey (2000) drei Empfehlungen für Usability Untersuchungen, um die Theorie mit der Praxis bei Verwendung von Ericsson & Simon (1993) Kategorien auszusöhnen:

- a) Daten nur aus beobachtbarer Leistung sammeln und keine verbalen Protokolle mehr erstellen. Dies beseitigt zwar die Diskrepanz, aber auch eine wertvolle Informationsquelle und berücksichtigt keine Ausnahmesituationen (Systemabsturz).
- b) Die Theorie nach Ericsson & Simon (1993) strikt anwenden, ohne Abweichungen (auch wenn es für Softwaretest unrealistische Vorgaben sind).
- c) Anwendung alternativer Theorien und Praktiken (bei Boren & Ramey wurde Sprachkommunikation verwendet).

Ericsson & Simon (1993) setzen voraus, dass die Formulierungsprozesse völlig von Kommunikation getrennt werden können. Durch diese Annahme beschreiben sie die ProbandInnen als völlig passive ZuhörerInnen, die zeitweise nicht existieren, sondern nur „laut denken“, als wären sie jeweils alleine im Raum. Laut Boren & Ramey (2000) ist dies nicht realistisch, da es immer zumindest ein(e) ZuhörerIn (TesterIn, ModeratorIn, AssistentInnen) und ein(e) SprecherIn gibt (meist die Testperson), abhängig davon wer gerade spricht und wie viel die ModeratorInnen intervenieren dürfen. Sie treten für eine freiere menschliche Kommunikation unter dem Vorbehalt, dass immer die Positionen gewahrt bleiben und von einer hierarchischen Perspektive der Forscher mit einer Testperson und nicht zwei ProbandInnen miteinander sprechen (Boren & Ramey, 2000).

Die wesentlichen Unterschiede zwischen Ericsson & Simon (1993) und Boren & Ramey (2000) werden in vier Ansätzen für die Verwendung von Sprachkommunikation (SprecherIn und ZuhörerIn) und „*usability testing*“ in Boren & Ramey (2000) beschrieben, wobei der Konflikt zwischen den beiden Theorien mit steigender Aufzählungsnummer zunimmt:

***(1) Die Rahmenbedingungen für die Verwendung von Sprachkommunikation helfen für einen produktiven Testdurchlauf.***

Boren & Ramey (2000) berufen sich auf den „Sprachgenres“ nach Bakhtin (n.d.), welcher der Meinung war, dass jede Sphäre in der Sprache verwendet wird, entwickelt ihre eigenen relativ stabilen Arten von Äußerungen. Diese Genres reflektieren die spezifischen Bedingungen und Ziele der menschlichen Aktivität in der sie vorkommen. Dies bedeutet, dass der Mensch verschiedene Sprachgenres anwendet, in Abhängigkeit von wer, was, wo, wann und warum kommuniziert (Boren & Ramey, 2000).

Die Sprachgenres bilden die Grundlage von Boren & Ramey (2000) neuen Ansatz und daraus abgeleitet sollte jeder Usability Test drei Gruppen von Rollen und Status definieren. Diese Rollen decken die Fragen der Sprachgenres Definition ab, nachdem Wer, Was und Warum:

1. *Das Produkt, welches getestet werden soll.* Die ProbandInnen des TimeRider Thinking Aloud Tests wurden, als Teil der Einschulung von Beginn an darauf hingewiesen, dass das zu untersuchende Objekt das Programm selbst ist und in keiner Art und Weise die Testperson. Wichtig ist herauszufinden, ob TimeRider den vorgesehenen Zweck für die Zielgruppe der UserInnen, welche das Programm bei der täglichen Arbeit dann verwenden sollen, erfüllt.
2. *Die ProbandInnen sind ExpertInnen und vorrangig SprecherIn/RednerIn.* Die Testpersonen dieser Magisterarbeit waren allesamt fertig ausgebildete ÄrztInnen, welche auch in der Humanmedizinbranche tätig waren. Diese ExpertInnen wurden gewählt, da sie eine repräsentative Abbildung der Anwender-Zielgruppe darstellten. Wichtig war das Verständnis der TeilnehmerInnen, dass ihre Erfahrung in der Humanmedizin benötigt wird und nicht die Erkenntnis als Amateur-Interface-KritikerIn. Es wurde Ihnen erklärt, dass ihre Expertise mit der sie die Aufgaben lösen der besondere Beitrag ist, der den Softwaredesignern fehlt.
3. *Der Usability Tester (ModeratorInnen & deren Teams) sind vorrangig lernend(e) SchülerIn und ZuhörerIn.* Der Test für diese Magisterarbeit selbst hat aber gezeigt, dass auch weitere „Rollen“ des Testteams übernommen werden müssen. Am Beginn bei der Einschulung der ProbandInnen ist der spätere Moderator der Gastgeber, welcher in die Umgebung und in alle zur Verfügung stehenden Hilfsmittel einführt, um danach bei der Einschulung in das Programm selbst, kurz zum Lehrer zu werden. Während des Tests übernahmen der Moderator und sein Team dann die Rolle der technischen ExpertInnen, um Hilfestellungen und technische Expertise anzubieten, da es sich bei dem zu testenden Programm um einen weiterentwickelten Prototypen handelte, indem auch Softwarefehler auftraten oder Unklarheiten für die ProbandInnen entstanden. Nach Boren & Ramey (2000) sollten diese nachrangigen Rollen jedoch eine untergeordnete Rolle spielen und weder im Detail vorgestellt noch von Bedeutung sein.

Diese Regeln sollte am Anfang des Test explizit definiert und vorgestellt werden, sowie während des gesamten Tests strikt durchgeführt werden. Jegliche Ablenkung oder „Herabstufung“ der ProbandInnen von der primären Rolle als ExpertInnen, welche die Aufmerksamkeit auf die ModeratorInnen lenkt, sollte verhindert werden (Boren & Ramey, 2000). Weitere Rollen die vor dem eigentlichen Test definiert werden müssen sind das wann und wo. Die Auswahl einer geeigneten Testumgebung, Equipment und Zeitpunkt (*timing*) sind

nur subtile dafür aber mächtige Faktoren, da sie die Wahrnehmung der eigenen Rolle der ProbandInnen wesentlich beeinflussen können. Ein steriles Testlabor mit Einwegspiegel, Videokameras und Tischmikrofone erzeugt eine andere Atmosphäre, als ein für diese Magisterarbeit verwendetes Besprechungszimmer im Institutsgebäude des Instituts für Gestaltungs- und Wirkungsforschung der Technischen Universität Wien. Zwei ProbandInnen ermöglichten es sogar an deren Arbeitsplatz den Softwaretest durchzuführen, dies erzeugte zusätzlich eine vertraute Arbeitsatmosphäre. Laut Boren & Ramey (2000) sind ebenfalls zu berücksichtigen: die Tageszeit zu der der Test stattfindet, die Länge der Zeit des Tests, die Ausstattung und Möbel im Raum, die Beleuchtung, das Computerequipment selbst und die Erfahrung der ProbandInnen mit Usability Tests, sowie alle Faktoren welche den Testpersonen helfen können in die Rolle der wissenschaftlichen ExpertInnen einzutauchen und sich nicht als Testobjekt selbst zu verstehen (Boren & Ramey, 2000).

*(2) Die Methode zur Datensammlung sollte die soziale Natur der Sprache berücksichtigen, aber nach der Ericsson & Simon (1993) Methode: mündliche Berichte erwirken, welche so viel wie möglich indirekt, ungestört und konstant als möglich sind.*

Nachdem die Rahmenbedingungen, wie im vorhergehenden Punkt beschrieben, umgesetzt wurden, sollten die Usability Tester sich an diese halten. Dadurch können die ProbandInnen unterstützt werden, die gestellten Testaufgaben mit einem Minimum an Ablenkung und Hilfe bearbeiten zu können, um dabei so konstant wie nur möglich die eigenen Aktionen und Geschehnisse in Worte zu fassen. Nach Ericsson & Simon (1993) sollte dies ausschließlich, dafür aber präzise und ausführlich nur am Beginn des Tests durchgeführt werden. Mit dem Start des Test ist gleichzeitig jede Kommunikation einzustellen, nur die Erinnerung der ProbandInnen daran weiter laut auszusprechen was sie gerade denken, wenn diese zu lange schweigen, ist zulässig. Unter den Gesichtspunkten Sprachkommunikation, kann die Unterbrechung von Stille durch kurze und emotionslose „Befehle“ als aggressive Form zwischenmenschlichen Kontakts gesehen werden. Ericsson & Simon (1993) sehen jedoch diese Kommunikationsform, als weder ablenkend, noch störend sondern als konstante Formulierung mit dem Ziel den ProbandInnen weitere Wörter zu entlocken. Boren & Ramey (2000) vertreten den dazu gegensätzlichen Standpunkt, dass mittels Sprachkommunikation ein menschlich natürlicherer Kommunikationsweg bessere Resultate liefern kann.

- *Kontinuierliche Verwendung von Bestätigungszeichen (acknowledgment tokens)*

In Boren & Ramey (2000) wird bei der Verwendung von Sprachkommunikation für Usability Tests, der Einsatz von Zeichen, im Sinne von sprachlichen Rückmeldungen, zur Bestätigung (in weiterer Folge als Bestätigungszeichen angeführt) empfohlen, z.B. „OK“, „Ja“ oder „Mm hm“. Ein vorsichtiger, nicht übermäßiger Einsatz dieser Bestätigungszeichen

kann den ZuhörerInnen (ModeratorInnen und deren Teams) dabei helfen, eine untergeordnete und nicht auffällige Reaktion zu geben, wenn dies seitens der ProbandInnen sprachlich angedeutet oder sogar verlangt wird (z.B. Ist ... so richtig? Soll ich...? Darf ich...?). Während des Usability Test dieser Magisterarbeit wurden Situationen bei fast jeder Testperson beobachtet, dass die jeweilig(e) SprecherIn (Testperson) bei vorhersehbaren Gedankenübergängen eine Reaktion des Zuhörers (Moderator) erwartete. Manche SprecherInnen gingen sogar soweit, dass sie den Blick vom Bildschirm abwendeten und den Moderator ansahen, bis dieser eine Reaktion zeigte und die jeweiligen SprecherInnen mit der Bestätigungsreaktion zufrieden waren. Es wurde beobachtet, wie von Boren & Ramey (2000) beschrieben, dass völlige Stille ohne jeglichen gesprochenen Laut des Moderators, die SprecherInnen dazu verführte die „Verbindung“ zum menschlichen „Gegenüber“ zu überprüfen, indem eine gezielte Frage an den Zuhörer gerichtet wurde. Im aktuellen Test konnte ebenfalls beobachtet werden, wie von Boren & Ramey (2000) beschrieben, dass der „Verbindungsüberprüfung“ entgegengewirkt werden konnte, indem während des Usability Tests der Moderator (Zuhörer) mittels kontinuierlichen aber unauffällig übermittelten Bestätigungszeichen, sich bemerkbar zu machen. Dadurch konnte sich die SprecherInnen mehr auf die Aufgaben konzentrieren und fließend und ungestört die Gedanken formulieren (Boren & Ramey, 2000).

- ***Die Wahl des Zeichens beeinflusst die mögliche Reaktion***

Boren & Ramey (2000) empfehlen aufgrund ihrer Literaturrecherche, die Verwendung der Bestätigungszeichen „mm hm“ und „uh-huh“ (= „mm hm“ mit offenen Mund) mit einer untersuchenden Intonierung. Diese beiden Bestätigungszeichen zwingen während eines Usability Tests die SprecherIn (Testperson) nicht zu(r) ZuhörerIn zu werden, da nicht der Eindruck entsteht, dass die ZuhörerInnen (ModeratorInnen und deren Team) diese Bestätigungszeichen gibt, weil plötzlich etwas gesagt werden will. Wenn diesen Bestätigungszeichen auch noch eine untersuchende Intonierung gegeben wird, dann werden die SprecherInnen gleichzeitig dazu eingeladen weiter zu machen und „laut zu denken“. Dies passiert jedoch ohne relevante Informationen zu den Aufgaben preiszugeben, ohne abzulenken oder weitere Kommentare der SprecherInnen einzufordern. Diese Bestätigungszeichen mögen den Eindruck vermitteln, dass sie nur dahin gesagt und zusammenhangslos sind, das Gegenteil ist jedoch der Fall, denn sie ermutigen die SprecherInnen weiter zu machen, da Ihre „Anfrage“ empfangen wurde und die ZuhörerInnen noch immer anwesend sind (Boren & Ramey, 2000).

- ***Frequenz von Bestätigungen***

Die Wiederholungsrate dieser Bestätigungszeichen ist schwer einheitlich zu definieren, da es auf mehrere Faktoren ankommt. Es sollte einerseits vorrangig auf die SprecherInnen (ProbandInnen) jeweils ankommen und wann diese anzeigen eine Bestätigung zu erwarten

(„Verbindungsüberprüfung“, Gedankenübergänge, direkte nachfragen). Andererseits sollten Bestätigungszeichen dem „Fluss“ der Kommunikation folgen und nicht forcierend sein und unterbrechend, sondern konstant und kontinuierlich einfließen. Diese Art von Kontakt über ein Headset mit Kameras zu führen, gestaltet sich vom Gesichtspunkt der menschlichen Kommunikation her schwierig. Der Test dieser Magisterarbeit wurde daher als eine der Varianten von Boren & Ramey (2000) durchgeführt, in der ein Moderator und sein Team sich im selben Raum wie die ProbandInnen befinden aber außerhalb deren Sichtweite saßen (Boren & Ramey, 2000).

- ***Argumente gegen kontinuierliche Bestätigungen aus Sicht der Sprachkommunikation***

Einerseits kann Stille bei einer Kommunikation zwischen Testperson (SprecherIn) und ModeratorIn (ZuhörerIn) als eine Art von Reaktion betrachtet werden, wenn auch als eine distanzierte und herablassende Art. Andererseits kann argumentiert werden, dass bei Bestätigungszeichen welche die SprecherIn empfängt, für dessen Verarbeitung kognitive Ressourcen benötigt werde, welche wiederum die Lösung der Aufgaben beeinflussen kann. Dies ist jedoch zu vernachlässigen, da auch im täglichen menschlichen Sprachgebrauch während gesprochen wird, die Verwendung von Bestätigungszeichen stattfindet ohne zwingend den Redefluss oder Gedankengang zu beeinträchtigen, daher kann dies auch für gültig in der Sprachkommunikation bei Evaluierungstests gesehen werden. Das letzte Argument kommt von Ericsson & Simon (1993), welche beim Verstummen der Testperson empfehlen, anstatt mit „Reden Sie weiter“, mit einen kontenfreien „Erinnerungston“ zu reagieren. Hier halten Boren & Ramey (2000) dagegen, indem Sie ihr Bestätigungszeichen „mm hm“ vorschlagen, da dieses einem einzelnen Tonsignal ähnlich ist aber natürlicher in die kontenfreie Kommunikation einfließt und weniger befremdlich wirkt, vor allem wenn es kontinuierlich eingesetzt wird (Boren & Ramey, 2000).

- ***Erinnerung der ProbandInnen weiter zu sprechen***

Wann eine Testperson inne hält und nicht mehr „laut denkt“, hängt nicht nur von der Person selbst und ihren verbalen Fähigkeiten ab, sondern auch von der Komplexität der Aufgabenstellungen und ob diese zum Ersten Mal durchgeführt werden, sowie ob das Programm vorher bekannt war oder nicht. Dieser Effekt trifft auch auf den Test dieser Magisterarbeit zu, da die ProbandInnen sowohl die Tasks als auch TimeRider zum Ersten Mal sahen. Erinnerungen sollten daher so unauffällig und nicht ablenkend sein, dass im besten Fall die Testpersonen sofort nach der Erinnerung weiter sprechen, ohne dass sie darüber nachdenken, dass Sie pausiert haben. Boren & Ramey (2000) empfehlen als Erstes ihr Bestätigungszeichen „mm hm“ zu verwenden, da dieses Zeichen am unverfänglichsten und ausreichend ist, die SprecherInnen daran zu erinnern, dass ZuhörerInnen warten. Darüber hinaus empfehlen sie die Verwendung der offeneren aber direkten und trotzdem konten-

freien Phrase: „Und jetzt?“ („*And now?*“). Es obliegt weiteren Tests und Untersuchungen andere einfache und kontenfreie Erinnerungen zu finden. Diese können jedoch nur dann als erfolgreich bezeichnet werden, wenn sie in einer sofortigen Wiederaufnahme des „laut Denkens“ resultieren, ohne Rückfragen, Entschuldigungen oder Erklärungen (Boren & Ramey, 2000).

**(3) *Interventionen seitens des Forschungsteams sind durch Eventualitäten welche beim Usability Test auftreten notwendig, wobei Ericsson & Simon (1993) keine spezifische Anleitung dafür geben, wie diese Interventionen unter Ihren Modell durchzuführen sind.***

Die Eventualitäten welche bei Usability Tests auftreten können, hängen einerseits von den Testpersonen und andererseits vom zu untersuchenden System selbst ab. Die zu testenden Systeme sind jedoch meist noch in Arbeit, welches unvorhergesehene Fehler und Systemabstürze zur Folge haben kann. Obwohl nicht zwingend das gesamte System oder dessen Interface getestet wird, kann fehlender Text oder Funktionalität, die ProbandInnen verwirren oder unerwartete Schwierigkeiten verursachen, da die noch nicht voll funktionsfähigen Teile meist nicht einfach separat vom Rest des Systems deaktiviert werden können. Das Testszenario ist eine Art Rollenspiel indem ModeratorInnen die ZuhörerInnen sind und die ProbandInnen die SprecherInnen sind, dadurch kann es zusätzlich zu unerwarteten Aktionen oder Fragen seitens der ProbandInnen kommen. Ericsson & Simon (1993) geben auf diese Eventualitäten keinerlei Hinweise oder Antworten, wie mit diesen umzugehen ist, da sie als kognitive Psychologen eine menschliche Testperson, deren Gedankengänge mittels eines Testsystems untersuchen und nicht das Testsystem selbst, wie es bei Usability Test der Fall ist (Boren & Ramey, 2000).

- ***Die Herausforderung bei anhaltenden Monologen seitens der ProbandInnen während des Usability Tests.***

In Ericsson & Simon (1993) wird nicht beschrieben wie mit einem Systemabsturz des zu untersuchenden Programms während des Usability Tests umzugehen ist. Boren & Ramey (2000) nehmen daher an, dass die folgenden zwei Punkte unter dem Ericsson & Simon (1993) System akzeptabel wären:

- a) Die ModeratorInnen lösen still das Problem und unterrichtet die Testperson darüber, dass es nach der Lösung weitergeht.***

Eine Lösung des Fehlers ohne jegliche Kommunikation wäre von einem sozialen Gesichtspunkt sehr unwahrscheinlich. Abgesehen davon könnten die UserInnen denselben Fehler wieder auslösen, da anzunehmen ist, dass dieser beim ersten Mal unbeabsichtigt ausgelöst wurde (Boren & Ramey, 2000).

b) *Die ModeratorInnen signalisieren eindeutig, dass der Test gestoppt wird.*

Nach dem offiziellen Stopp des Tests, können die ModeratorInnen oder jemand aus deren Team, den Fehler beheben und danach die ProbandInnen instruieren wie weiter zu verfahren ist. Danach kann offiziell wieder weitergetestet werden und die Testperson startet wieder mit Thinking Aloud (Boren & Ramey, 2000).

- *Das System stürzt ab, ein schwerwiegender Fehler tritt auf oder der Prototyp ist nicht vollständig.*

Aus Sicht der Sprachkommunikation nach Boren & Ramey (2000), ist es wichtig trotz der besonderen Umstände, die Rollenverteilung vor dem Fehler beizubehalten (ModeratorIn = SchülerIn; ProbandIn (Testperson) = LehrerIn und FachexpertIn; TimeRider = Testobjekt). Die Testperson sollte nochmals darauf hingewiesen werden, dass es nicht deren Schuld ist, dass der Fehler aufgetreten ist, sondern die Schuld des Programms. Sollte der Fehler noch nicht bekannt sein, können ModeratorInnen der Testperson dies als Beweis derer Fachexpertise aufzeigen. Dieser Hinweis sollte nicht übertrieben dargestellt werden und ist je nach Auftreten der Testperson unterschiedlich zu vermitteln oder gar nicht erst zu erwähnen. Manche ProbandInnen brauchen jedoch diese Art der Bestätigung, da diese sonst sich verunsichert fühlen oder im schlimmsten Fall sich selbst die Schuld an dem Fehler geben: „...ich verschwende nur Ihre Zeit!“ oder „Jetzt habe ich es kaputt gemacht!“. Je nach Schwere des Fehlers ist von den ModeratorInnen zu entscheiden, ob eine kurze Pause eingelegt werden sollte oder ob der Fehler einfach und schnell zu beheben ist (Boren & Ramey, 2000).

- *Die Testperson übersieht oder vermeidet wichtige Funktionalität.*

Die ProbandInnen könnten bei der Lösung der Aufgabe wichtige Teile der Aufgabenstellung oder des zu untersuchenden Teils des Programms unbeabsichtigt umgehen, oder auf eine andere Art als vorhergesehen lösen. Dies ist bei Usability Tests ein Problem, da die Untersuchung und Datensammlung von Schlüsselstellen und Schlüsselaktionen, das Ziel der Tests ist. Wenn diese Bereiche nicht erfasst werden, wäre der Test als solches überflüssig, da keine Erkenntnis daraus gewonnen werden kann. Nach Boren & Ramey (2000) bietet Sprachkommunikation eine Lösung an, indem die Rollen der Kommunikationspartner getauscht werden können (Boren & Ramey, 2000).

Durch die Anwendung von Sprachkommunikation bei der Lösung von unerwarteten Eventualitäten, ist am wichtigsten, dass die Testperson sich nie erniedrigt, peinlich berührt oder törricht fühlt. Sollte die Testperson die Anweisungen der Aufgabenstellung nicht verstanden haben, dann sollte behutsam vorgeschlagen werden die Aufgabenstellung nochmals durchzulesen. Sollten die ProbandInnen aber denken, dass die Aufgabe erfüllt ist, obwohl

sie es aber tatsächlich noch nicht ist, dann schlagen Boren & Ramey (2000) vor, dass auf jeden Fall interveniert werden muss, da ansonsten kritische Programmfunktionen ungetestet bleiben. Da es keine weiteren Erläuterungen dazu gab, außer dass weitere Untersuchung zu diesem Fall durchgeführt werden sollten, entschloss sich das Forschungsteam dieser Magisterarbeit die ProbandInnen, bei denen dies auftrat, vorsichtig nachzufragen, ob die Aufgabenstellung zur Gänze erfüllt sei und ob gegebenenfalls eine andere Interpretationsmöglichkeit der Aufgabenbeschreibung möglich ist. Diese Herangehensweise wurde gegenüber den ProbandInnen damit argumentiert, dass das Forschungsteam aus KEINEN Experten in Sachen Humanmedizin besteht und man als Forscher mit Informatik Hintergrund, danach fragt ob alle logischen Möglichkeiten ausgeschöpft sind. Dadurch wurde die Rolle der ProbandInnen (ExpertInnen) und des Forschungsteams (Zuhörer) beibehalten und explizit nach deren Expertise gefragt. Diese Vorgehensweise wurde ebenfalls für den Fall angewandt, falls die ProbandInnen von der zu überprüfenden Funktionalität abweichen und aus deren Kreativität heraus oder aufgrund von Systemfehlfunktionen, die Aufgabe auf eine andere Weise als erwartet lösten (Boren & Ramey, 2000).

Ein Problem welches auch während der Tests dieser Magisterarbeit auftrat war, dass die ProbandInnen nicht weiter wussten oder gedanklich nicht mehr weiter wussten. In einen solchen Fall weisen Boren & Ramey (2000) darauf hin, dass es entscheidend ist solch eine Situation zu entschärfen, da die ProbandInnen schnell frustriert oder verlegen sind und sich „dumm“ oder unzulänglich fühlen könnten. Diese negative Stimmung kann den weiteren Testverlauf beeinflussen, vor allem wenn eine derartige Situation nochmals im späteren Verlauf eintritt. Es wird empfohlen als erstes sich dankbar zu zeigen und zu bestätigen, dass ERFOLGREICH ein neuer Fehler, aufgrund der Expertise der ProbandInnen gefunden wurde und wie wertvoll dies für die durchgeführte Studie ist. Jegliche Schuld sollte auf das Programminterface geschoben werden und nochmals die UNSCHULD der ProbandInnen betont werden. In den Tests dieser Magisterarbeit wurde dies umgesetzt, indem sowohl in der Einschulungsphase, als auch während der Test (je nach Situation entsprechend) die ProbandInnen darauf hingewiesen wurden, dass das Ziel der Tests sei, die Programmfehler zu entdecken welche der Usability schaden. Dadurch konnte das zusätzliche Risiko minimiert werden, dass ProbandInnen sich trotzdem schuldig fühlen und das Gefühl haben, dass die ModeratorInnen und das Forscherteam die Schuld nur versuchen herunter zu spielen. Somit bekamen die ProbandInnen ein Gefühl vermittelt, dass sie die Hauptaufgabe erfüllen, wenn Fehler oder Unklarheiten auftreten (Boren & Ramey, 2000).

Es werden in Boren & Ramey (2000) auch weitere Eventualitäten angeführt, welche Kommunikation erfordern und da sie beobachtet wurden, jedoch ohne Konkrete Lösungsansätze und mit dem Verweis auf dafür weitere durchzuführende Untersuchungen. Ein Beispiel ist, wenn die ProbandInnen eine Frage stellen, welche sich durch die Aufgabenstellung erklären lässt. Diese Herausforderung wurde in dieser Magisterarbeit dadurch ge-

löst, indem, wie vorhin beschrieben, vorsichtig nachgefragt wurde, ob die Aufgabenstellung zur Gänze erfüllt sei und ob gegebenenfalls eine andere Interpretationsmöglichkeit der Aufgabenbeschreibung möglich ist. Diese Lösungsstrategie wurde in dieser Magisterarbeit ebenfalls für den Fall angewendet, wenn die ProbandInnen Fragen über die Aufgabe welche soeben bearbeitet wird stellten. Dadurch wurde signalisiert, dass die Aufgabenstellung alle relevanten Informationen enthielten. Sollte eine ProbandInnen sich mit dieser Antwort nicht zufrieden geben, dann wurde darauf hingewiesen, dass das Programminterface die BenutzerInnen zur Antwort leiten sollte und wenn es dies nicht kann, handelte es sich um einen neu entdeckten Fehler.

Es kam aber auch in den Tests dieser Magisterarbeit vor, dass die ProbandInnen eine Frage stellten, aus der sich erkennen ließ, dass die Aufgabe mittels einer anderen Funktionalität bearbeitet werden wird als vorgesehen. Der Moderator entgegnete darauf hin, wie von Boren & Ramey (2000) empfohlen: „Was würden Sie sich vom Programm erwarten?“ und bestärkte die Testperson weiter zu machen mit dem sanften Hinweis, nochmals die Aufgabenstellung zu lesen und im hartnäckigsten Fall, ob es vielleicht noch weitere Lösungswege gibt.

Ebenfalls neigten einige ProbandInnen während der Tests dieser Magisterarbeit zu kommunikativ zu werden. Dies erkannte man an länger gehaltenen Monologen, welche im Sinne der Aufgabenstellung nicht relevant waren oder durch den Versuch mit dem Moderator in eine Art Gespräch zu beginnen. In beiden Fällen wurde je nach Person entschieden, wann darauf sanft hingewiesen wurde, wieder anhand der Aufgabe weiter zu arbeiten. Wie von Boren & Ramey (2000) empfohlen, wurde hier, dass Bestätigungszeichen „OK“ verwendet: „OK, wo waren wir stehengeblieben“. Beim Versuch ein Gespräch zu beginnen wurde, darauf hingewiesen, dass für alle Test relevanten Fragen am besten die Aufgabenstellung zu konsultieren ist oder gerne nach dem Test genügend Zeit für weitere Gespräche vorhanden ist.

***(4) Proaktive Interventionen welche Informationen den ProbandInnen entlocken, aber nicht erlaubt sind nachdem Modell von Ericsson & Simon (1993).***

Bisher wurden zwei Formen von Sprachkontakten beschreiben, einerseits die Bestätigungszeichen, welche kontinuierlich und parallel zum Thinking Aloud Test von den ModeratorInnen eingestreut werden sollten, und andererseits der erforderliche Kontakt bei unvorhersehbaren Ereignissen. Erster ist auch im Modell von Ericsson & Simon (1993) vorgesehen, der zweite wird nicht im Detail beschrieben, da nach Beginn des Test vor allem absolute Stille im Testraum herrschen muss. Boren & Ramey (2000) stellen auch eine dritte Form des Sprachkontakts vor, indem Usability Tester in einer Art und Weise in Kontakt mit den ProbandInnen treten, die weder in das Bild eines ungestörten Testablaufs passen, noch aus Testsituationen heraus entstehen. Diese dritte Form des Sprachkontakts, enthält

alle Interventionen seitens des Forschungsteam zu den jeweiligen ProbandInnen. Dies kann aufgrund verschiedenster Motivationen basieren, aber immer mit demselben Ergebnis, an zusätzliche Informationen für die Usability Diagnose zu kommen, welche aber nicht durch das Modell von Ericsson & Simon (1993) gesammelt werden könnten. Boren & Ramey (2000) unterscheiden dabei zwei Motivationen, wann ein Usability Tester an zusätzliche Informationen kommen kann: Einerseits wenn die ProbandInnen unklare Kommentare geben und andererseits wenn das Forschungsteam mehr Informationen braucht, als bisher gesammelt werden konnten (Boren & Ramey, 2000).

***a. Unklare Kommentare der ProbandInnen ansprechen.***

In dieser Kategorie, geht es jedoch nicht darum, dass die ProbandInnen undeutliche Kommentare abgeben, dies kann immer wieder vorkommen und sollte aufgrund des Denk- und Redeflusses aber wenn möglich seitens der ModeratorInnen oder des Forschungsteams nicht unterbrochen werden. Es geht um Situationen in denen der unverständliche Kommentar möglicherweise essentiell sein könnte, wie z.B. eine besondere Funktion oder aus vorherigen Tests bekannte schwierige Stelle. Die Aufgabenstellung könnte auch kurz vor der Lösung stehen und die Tester wollen aber noch ein bestimmtes Detail oder einen Fehler untersuchen, bevor die ProbandInnen gedanklich den Task abschließen, denn das könnte wieder als eigener Fehler der Testperson interpretiert werden. Dasselbe Risiko besteht, wenn ModeratorInnen oder eine Person aus deren Team direkt nachdem unklaren Kommentaren nachfragen würde, daher gilt auch für diese Situation: behutsames Vorgehen. Ebenfalls muss bevor eine Intervention stattfindet bedacht werden, ob sich diese lohnt, da die Aufmerksamkeit der Testperson auf den Fragesteller gerichtet wird und der Fokus nicht mehr auf das Programm gerichtet ist. In weiterer Folge könnte die Testperson verunsichert werden, wodurch der weitere Testverlauf beeinflusst werden könnte (Schuldempfinden). Die in Boren & Ramey (2000, p.275) zusätzlich empfohlene Variante der Videoanalyse des Verhaltens der Testperson durch die Testperson selbst, wurde im Zuge der Test dieser Magisterarbeit nicht durchgeführt, da dies mit der spärlichen Zeit die den ÄrztInnen zur Verfügung stand, unvereinbar gewesen wäre.

Boren & Ramey (2000) empfehlen die Verwendung von kurzen und nicht in Befehlsform gestellten Fragen. Sie stellen den Anspruch dass ein einziges Wort sich am besten eignen würde. Welches Wort, hängt jedoch direkt von dem unklaren Kommentar der Testperson ab. Wenn ein Kommentar ausgesprochen wird wie: „Das ist komisch?“ dann sollten ModeratorInnen nachfragen „Komisch?“ oder wenn gefragt wird „Das ist interessant?“ dann sollte nachgefragt werden: „Interessant?“, und so weiter. Der Vorteil dieser minimalistischen Frage ist, dass sie den natürlichen Regeln der Sprache entspricht und sogar aus Sicht von Ericsson & Simon (1993) zulässig wäre (Boren & Ramey, 2000).

### ***b. Mehr Informationen herausholen.***

Dieser letzte Punkt entspricht nicht in dem Modell nach Ericsson & Simon (1993) der Informationsklasse *Level 3*, da das direkte Fragen nach Informationen mit einer größeren Wahrscheinlichkeit die Performance der Aufgabenlösung beeinflussen wird, da sie über die Abfrage im Kurzzeitgedächtnis hinaus geht und das Langzeitgedächtnis hinzugezogen werden muss. Boren & Ramey (2000) empfehlen daher, wenn es sich um eine fundamentale Frage bezüglich der Gedankengänge der Testperson handelt, sollte darauf verzichtet werden, da nach Ericsson & Simon (1993) die ProbandInnen nicht zuverlässig und unvoreingenommen über Ihre eigenen Gedanken urteilen können, um daraus ein kognitives Modell erstellen zu können. Handelt es sich jedoch um Fragen, um mehr Informationen über das Design des Programms zu erfahren, können diese Fragen wertvolle Einblicke in die Meinungen und Ansichten der Testpersonen liefern, welche wiederum wichtige Daten für die Verbesserung des Programms liefern können. Boren & Ramey (2000) weisen zum Abschluss jedoch darauf hin, dass weder Ericsson & Simon (1993) noch irgendein anderes Thinking Aloud Protokoll, spezifische Anweisung oder Strategien für eine solche Befragung bereitstellt (Boren & Ramey, 2000).

Diese Magisterarbeit ist dafür ausgelegt, ein neuartiges medizinisches Programm zu optimieren. Wie in diesem Kapitel detailliert beschrieben, ist auch in dieser Arbeit die Hauptmotivation, die Entdeckung aller Usability Probleme in der TimeRider Version 3.4. Dafür wurde der Ansatz von Boren & Ramey (2000) gewählt, da dieser Ansatz eine freiere Gestaltung der Thinking Aloud Tests auf Basis der Sprachkommunikation zulässt und somit ein gezieltes Suchen nach Usability Fehlern ermöglicht.

## 7 Didaktische Verwendungsmöglichkeit

Die Verwendung des Programms TimeRider in der Humanmedizin und deren didaktische Nutzbarkeit in der humanmedizinischen Ausbildung wurden neben der Usability-Fehleranalyse und der Interaktionsanalyse des Programms erhoben. Während der Untersuchung von TimeRider vermittelten die ProbandInnen den subjektiven Eindruck, dass TimeRider auch für die humanmedizinische Lehre eindeutig geeignet ist. Dies wurde immer wieder seitens der ProbandInnen betont und jeder der ProbandInnen hat begeistert empfohlen, dass untersuchte Programm ebenfalls in der Ausbildung von MedizinerInnen zu verwenden. Nach Ansicht der ÄrztInnen eignet sich TimeRider sowohl zum Training für Humanmediziner selbst, als auch für die Ausbildung während des humanmedizinischen Studiums und der StudentInnen. Im Fokus stand dabei vor allem die verwendete Visualisierungsmethode der Animation (siehe Kapitel 4). Die ProbandInnen betonten, dass die Verwendung von Animation im Gegensatz zu statischen Darstellungen, sowohl motivierend als auch hilfreich ist, wenn man sich einen Überblick über die Daten und deren Entwicklung über die Zeit schnell verschaffen will und Trends erkennen will.

In der Literatur finden sich dazu mehrere Standpunkte bezüglich Animation in der humanmedizinischen Lehre. Cingi (2013) führte eine Untersuchung durch um nachzuweisen, dass die Verwendung von Animation in der theoretischen humanmedizinischen Lehre einen Vorteil gegenüber statischen Darstellungen hat. Dabei wurden 20 ProbandInnen mittels Animation eingeschult und 20 ProbandInnen auf herkömmliche Art und Weise, mittels theoretischer Information und einfachen statischen Darstellungen. Der Lernerfolg wurde nach einer Lernphase, durch mündliche Tests überprüft. Nach dieser Überprüfung wurde jenen Testpersonen, welche nur statische Bilder zur Verfügung hatten, auch das Material mit der Animationsfunktion zugänglich gemacht, die ProbandInnen sollten beide Unterrichtsmaterialien nochmals verwenden und vergleichen. Diese Gruppe wurde dann nochmals mittels desselben mündlichen Tests geprüft, um den Einfluss durch das Animations-Lernmaterial zu dokumentieren. Die Ergebnisse der drei mündlichen Tests waren eindeutig und zeigten einen höheren Lerneffekt nach der Schulung mit Animationen. Die Gruppe welche von Beginn an mittels Animationen geschult wurden, konnten die Testfragen alle detaillierter beantworten, als die zweite Gruppe, welche mit statischen Bildern und Texten

lernen konnten. Nachdem die zweite Gruppe den Zugang zu den Materialien mit den Animationen hatte, wurde auch im zweiten Test beobachtet, dass ebenfalls ein detaillierteres Wissen vorhanden war. Daraus schließt Cingi (2013), dass Animationen in der humanmedizinischen Ausbildung hilfreich sind. Durch den Vergleich der Testergebnisse, welche auf zwei verschiedenen Schulungsmaterialien basierten, konnte beobachtet werden, dass Animation ein effektiveres Lehrinstrument ist (Cingi, 2013).

In Ruiz, et al. (2009) wurde eine kritische Literaturanalyse hinsichtlich Computeranimation in humanmedizinischer Ausbildung durchgeführt. Darin wird Animation als Darstellung von dynamischen Veränderungen über die Zeit und Raum beschrieben, welche Phänomene und Konzepte veranschaulichen kann, die sonst schwierig darzustellen wären. Es wird auch darauf hingewiesen, dass Animationen nicht immer effektiv sein müssen, denn es sind die Prinzipien der Animation zu beachten, die deren Verwendung bestimmen. Eine Literaturrecherche wurde durchgeführt, um das Potenzial für Applikationen mit Animationen in der humanmedizinischen Lehre aufzuzeigen und um evidenzbasierte Prinzipien für deren Design und Verwendung zu identifizieren. Die meisten Untersuchungen welche ausgewertet wurden, beschrieben einen Vergleich zwischen Animationen und alternativen computerbasierten oder nicht computerbasierten Darstellungen. Die Ergebnisse zeigten, dass in der humanmedizinischen Lehre eine Vielzahl von computerunterstützten Applikationen verwendet wird, dabei aber nur wenige Vergleichsstudien veröffentlicht wurden und die Beweislage nicht schlüssig ist. In Forschungsarbeiten außerhalb der humanmedizinischen Lehre, in denen Animationen und statischen Bildern verglichen werden, werden widersprüchliche Ergebnisse beschrieben. Dies ist ein möglicher Hinweis auf eine erhöhte kognitive Auslastung durch Animationen oder der illustrierten Bewegung. Die Vorteile von Animationen sind abhängig von den Charakteristiken des Lernenden, wie Vorwissen und räumliche Fähigkeiten. Animationsfunktionen welche den Lernprozess unterstützen sollen, enthalten meist die Kontrolle der UserInnen über die Animationsschritte, welche es erlauben mit der Animation zu interagieren und somit den Animationsvorgang in kleinere Informationsbrocken (= *information chunks*) aufzuteilen (Ruiz, et al., 2009).

Ruiz, et al. (2009) empfehlen daher weitere Forschungen über Animationen in der humanmedizinischen Lehre vorzunehmen, da die bereits vorhandenen Forschungsarbeiten zu wenig Informationen über die Lerneffekte durch Verwendung von Animationen enthalten. Weitere Untersuchungen auch für und in anderen Forschungsbereichen sind essentiell um zu ermitteln, wann und wie die Animationen am effektivsten eingesetzt werden können (Ruiz, et al., 2009).

## **8 Evaluierung von TimeRider**

In diesem Kapitel 8 werden die Ergebnisse der Untersuchung von TimeRider präsentiert. Im nachfolgenden Unterkapitel 8.1 wird die Untersuchung, deren Aufbau und Rahmenbedingungen beschrieben. Darauf folgt unter 8.2 eine kurze Einführung in das Programm TimeRider. Das Unterkapitel 8.3 enthält Informationen zu den ProbandInnen und klärt die Frage warum 10 Testpersonen für diese Magisterarbeit ausreichend sind. Das darauffolgende Unterkapitel 8.4 enthält die zentrale Fragestellung und die daraus entstandenen Hypothesen. Im letzten Kapitel 8.5 werden die Resultate präsentiert.

### **8.1 Beschreibung der Untersuchung**

Das Ziel der für diese Magisterarbeit durchgeführten Untersuchung von TimeRider, war es anhand der Ergebnisse der Testdurchläufe, einen Aufschluss über die gravierendsten Utility und Usability Mängel des Programms zu erlangen. Im Zuge der Untersuchung wurden 10 ProbandInnen jeweils einzeln getestet. Alle ProbandInnen waren ÄrztInnen (im Gegensatz zu vielen Studien, welche nur StudentInnen verwendeten, vgl.: Pohl, et al., 2012; Martins, et al., 2008) aus verschiedenen medizinischen Fachbereichen, wobei vier davon Frauen und die restlichen sechs Männer waren. Das durchschnittliche Alter mit einer Ausnahme lag zwischen 26 und 35, denn eine einzelne Person war um die 50. Die ÄrztInnen wurden erst nach der Erstellung des Programms hinzugezogen und sahen daher die Software zum ersten Mal.

Bei acht von zehn Testdurchgängen konnte ein ruhiger Raum, dank Frau A.o.Univ.-Prof.in Dr.in Mag.a Margit Pohl in den Räumlichkeiten des Institutsgebäudes des Instituts für Gestaltungs- und Wirkungsforschung an der Technischen Universität Wien benützt werden. Bei den zwei anderen Tests wurde auf das Angebot der beiden ÄrztInnen eingegangen, aus terminlichen Gründen am jeweiligen Arbeitsplatz der ProbandInnen die Tests durchzuführen. Dabei wurde darauf geachtet die räumlichen Bedingungen des Instituts der anderen acht Tests nachzubilden, damit eine einheitliche Testumgebung gewährleistet werden konnte.

Getestet wurde die „Programmversion 3.4“ des Programms TimeRider, auf Utility und Usability Mängel. Die während der Untersuchung durchgeführten Aktionen der ProbandInnen am Bildschirm, der Bildschirminhalt, die Bewegungen des Mauszeigers und die getätigten Aussagen, ausgesprochenen Gedankengänge und Kommentare wurden mit dem Bildschirmaufzeichnungsprogramm „Microsoft Windows Media Encoder 9“ aufgezeichnet. Dieses Programm speichert die Aktionen am Bildschirm und die Aussagen der ÄrztInnen als „Windows Media Audio/Video file“ mit der Dateiendung „WMV“ ab. Die Verwendung dieser Software wurde einer Videokamera Aufnahme vorgezogen, da der Bildschirminhalt deutlicher zu sehen ist und andernfalls wichtige Details verloren gegangen wären. Nach einer Einführung und von Beginn der Untersuchung an hin zum Ende der letzten Aufgabe wurde ohne Pausen aufgezeichnet. Das anschließende kurze Interview wurde jedoch separat mit einem analogen Diktiergerät aufgezeichnet (siehe Fels, 2015 für Details).

Der Versuchsaufbau bestand aus einem mobilen Computerlaptop der Marke *HP Elitebook Compaq 6930p* auf dem das Betriebssystem Windows Vista 64bit Ultimate und Java 6 (*jre6 update 16*) installiert war. Weitere technische Bestandteile der Untersuchung waren ein über VGA angeschlossener Samsung 20 Zoll TFT Flachbildschirm mit einer maximalen Auflösung von 1680 x 1050 Pixel, eine optische drei-Tasten-USB-Computermaus mit Scrollrad und ein kabelgebundenes 3,5 mm Klinke Mikrofon, welches um den Hals der jeweiligen ProbandInnen gelegt wurde und vorne auf der Brust ruhte. Die ProbandInnen saßen auf einen Stuhl und nur der Bildschirm und die Computermaus wurden auf einem Tisch in Reichweite direkt davor platziert. Die Computermaus ermöglicht die vollständige Bedienung des Programms und der Laptop mit eingebauter Tastatur wurde neben dem Bildschirm platziert, damit die ProbandInnen bei Wunsch die Tastatur ausprobieren oder verwenden konnten. Für die Darstellung der Windows Desktop Oberfläche wurde ein einheitliches neutrales Grau gewählt, damit keinerlei Beeinträchtigung davon ausgehen konnte.

Die Patientendaten der Kohortenstudie wurden von Untersuchungen und Nachuntersuchungen in einer Diabetes-Ambulanz inspiriert. Als Diabetes mellitus werden chronische Stoffwechselerkrankungen bezeichnet, bei denen zu wenig Insulin gebildet wird. Insulin ist ein Hormon und wird in der Bauchspeicheldrüse (Pankreas) gebildet, es regelt den Blutzuckerhaushalt in dem es den Blutzuckerspiegel senkt. Jeder unbehandelte Patient hat erhöhte Blutzuckerwerte (Hyperglykämie) und hat ein erhöhtes Risiko Komplikationen ausgesetzt zu sein, wie z.B. dem diabetischen Koma oder Herzkreislauferkrankungen. Die Behandlung der Krankheit hat große Veränderungen im gesamten Lebensstil des Betroffenen zur Folge, angefangen von oraler Medikation über zusätzlich oder stattdessen intravenös verabreichten Insulin mittels Spritzen, bis hin zu einer speziellen Diät. Die Wahl der Therapie ist dabei abhängig von mehreren Faktoren, wie z.B. vom Diabetes Typ und von Begleiter-

krankungen, aber auch von der Erfahrung sowie der Bereitschaft des Patienten die Therapie gewissenhaft durchzuführen.

Der fiktive Datensatz sollte aussagekräftige Analysen und Diagnosen zulassen und daher wurden ExpertInnen hinzugezogen, welche über das notwendige Fachwissen verfügten um daraus einen repräsentativen Datensatz erstellen zu können. Die Untersuchungsintervalle sollten zwischen sechs Wochen bis drei Monaten sein und die Daten sollten unregelmäßig erfasst sein. Die Inkonsistenz der Daten spiegeln den medizinischen Alltag wieder, da die Messungen immer abhängig von der Verfassung des einzelnen Patienten sind. Es wurde eine Einigung erzielt, dass die Daten aus 20 fiktiven Patienten, 10 quantitativen Variablen (z.B. Nüchtern-Blutzucker = NBZ) und 22 binären Variablen (z.B. Insulintherapie) bestehen sollten.

Die Untersuchung von TimeRider selbst bestand aus drei Phasen. Die erste Phase war eine kurze Einführung in die Forschungsthematik, den Ablauf und in die Benutzung des Programms. Die ProbandInnen sollten einen geistigen Orientierungspunkt bekommen, warum dieses Programm untersucht wird, wonach geforscht wird und welche Messungsmethode angewandt wird. Den Testpersonen war weder die Funktionsweise der Untersuchungsmethode von Thinking Aloud (siehe Kapitel 6.1 für Details), noch das Programm TimeRider bekannt. Daher wurde jede Testperson sowohl in Thinking Aloud, als auch in die Funktionsweise des zu untersuchenden Programm-User Interface eingewiesen. Hierfür wurde ein Dokument „TimeRider Einschulung“ (siehe Anhang A.1 für Details) den ProbandInnen jeweils ausgehändigt, wo die Rahmenbedingungen grob umrissen und erklärt wurden. Jeder Testperson wurde das User Interface von TimeRider mit einheitlicher Wortwahl und Abläufen erklärt. Die Einführung wurde vom Autor dieser Magisterarbeit selbst und Herrn Ulrich Fels, BSc durchgeführt. Diese erste Phase dauerte ca. 15 Minuten für jeden der ProbandInnen.

Die zweite Phase bildeten die insgesamt vier Aufgabenstellungen (siehe Anhang A.4 für Details) an die ProbandInnen, welche zu lösen waren. Alle vier Tasks welche unabhängig voneinander von den ProbandInnen zu lösen waren, wurden mit Hilfe von ÄrztInnen erstellt, welche das benötigte Fachwissen hatten.

Vor Beginn der Tasks wurde den Testpersonen vermittelt, dass weder ihr Wissensstand noch ihre Fähigkeiten, vor allem im Hinblick auf ihren Fertigkeiten im Umgang mit Computern, getestet oder geprüft werden, sondern nur das Programm TimeRider mit seinen Features, seiner Benutzungsfreundlichkeit und seinen Fehlern. Im Fokus stand nicht wie die ProbandInnen gehandelt haben oder ob eine Aktion, Aussage oder Idee positiv oder negativ war, sondern warum das Programm sie zu dieser Handlung geführt hatte. Bei bestimmten Aktionen können Unsicherheit oder Unverständnis auftreten, dies dürfen die BenutzerInnen jedoch nicht als eigenen Fehler auffassen. Der Fehler oder das Versagen liegt

immer beim Programm selbst. Die ProbandInnen, die darauf nicht hingewiesen werden, könnten sich selbst die Schuld dafür geben und nur wenige Gedankengänge und negative Kritik welche dann analysiert werden könnten, würden überhaupt ausgesprochen werden.

In der ersten der drei Aufgabenstellungen wurden konkrete Parameter für die X und Y-Achse vorgegeben, damit die ProbandInnen sich mit der für sie unbekanntem Software vertraut machen konnten. Im Detail gab der erste Task folgende Achsenparameter vor: „x-Achse: BMI, y-Achse: HbA1c [%]“. Die eigentliche Aufgabenstellung war: „Experimentieren Sie nach Belieben mit dem Programm und machen Sie sich vertraut mit der Oberfläche und den Funktionen. Beschreiben Sie ihre Erkenntnisse“.

Die Aufgaben zwei und drei stellten dann konkrete Anforderungen und Fragen welche zu erfüllen waren. In Task II war die Vorgabe: „x-Achse: BMI, y-Achse: RR syst [mmHG]“. Die Aufgabenstellung war: „Heben Sie das Geschlecht der Patienten hervor (z.B. mittels Farbe oder Form)“. Des Weiteren wurde gefragt: „Zum letzten Untersuchungszeitpunkt könnte man die Patienten in zwei Gruppen einteilen, abhängig von Ihrem Gesundheitszustand. Welche Patienten haben gute Werte?“. Damit die ProbandInnen auch aktiv Interaktionen durchführten, wurde verlangt: „Wählen Sie jeweils eine typische (durchschnittliche) Person aus den zwei Punktwolken aus. Begründen Sie ihre Auswahl.“. Im dritten Task war die vordefinierte Vorgabe: „x-Achse: NBZ, y-Achse: RR diast [mmHG]“. Die konkrete Aufgabenstellung war: „Schränken Sie den angezeigten Datensatz wie folgt ein:  $NBZ \leq 100$ ;  $RR \text{ diast.} \leq 80$ “. Die ProbandInnen wurden angehalten mit den verschiedenen Interaktionsmöglichkeiten zu experimentieren: „Wählen Sie eine Darstellung, die einen guten Überblick über den Verlauf aller verbleibenden Patienten ermöglicht.“ Eine weitere Angabe war: „Bei welchen Personen ist eine günstige Entwicklung zu erkennen? Wie entwickelt sich der allgemeine Gruppentrend?“. Nach den ersten Einstellungen sollte wie folgt vorgegangen werden: „Experimentieren Sie nach Belieben. Beschreiben Sie ihre Erkenntnisse.“

Die vierte und letzte Aufgabe ermöglichte es den Testpersonen frei zu experimentieren und die Software nach eigenen Vorstellungen zu verwenden (siehe Anhang A.4 für Details). Die Vorgabe war: „Keine“ und die Aufgabenstellung war: „Nutzen Sie gewonnene Eindrücke und Erkenntnisse, um beliebig mit dem Programm zu experimentieren und gegebenenfalls Unklarheiten zu beseitigen oder Interessantes genauer auszuprobieren. Beschreiben Sie ihre Erkenntnisse.“.

Ein Dokument indem die vier Aufgabenstellungen, welche von jeder Testperson eigenständig gelöst werden sollten, wurden jeder Testperson in ausgedruckter Form übergeben (siehe Anhang A.4 für Details). Ebenfalls wurde jeder Testperson eine Tabelle (siehe Anhang A.3 für Details) mit den im Programm verwendeten Abkürzungen (z.B. RR diast = diastolischer Blutdruck) ausgehändigt. Diese Abkürzungen sind gängige Abkürzungen in der Humanmedizin für medizinische Werte. Obwohl die ÄrztInnen aus verschiedenen me-

dizinischen Fachbereichen stammten, waren sie aufgrund der gemeinsamen medizinischen Grundausbildung, darauf geschult die Korrelation zwischen bestimmten Parametern zu erkennen (z.B. Übergewicht und Blutzuckerspiegel). Die Liste diente nur als Hilfestellung, wenn gewisse Begriffe nicht geläufig waren. Darüber hinaus wurde eine Liste mit den wichtigsten Parameterdefinitionen ausgehändigt (siehe Anhang A.2 für Details). Beide Listen konnten je nach Bedarf herangezogen werden und dienten als Nachschlagewerk und standen jederzeit zur Verfügung.

Die Aufgaben waren darauf ausgelegt das Programm und seine Funktionen zu erforschen, daher gab es auch keine vordefinierten Antworten auf die Aufgabenstellungen. Die allgemeine Gültigkeit der entdeckten Mängel beruht darauf, dass nicht die Erfüllung der Aufgabenstellungen bewertet wurde, sondern wieso es zu einem Empfinden, einer Aussage, einer Aktion der ProbandInnen kam. Sollte beispielsweise eine Testperson eine fehlende Funktion bemängeln, obwohl sie im Programm existiert aber nicht gesehen oder gefunden wurde, dann können daraus Rückschlüsse auf Fehler im Programmdesign gezogen werden. Bei einem Programmdesign indem die Testperson die gewünschte Funktion findet, würde hingegen kein Fehler vermerkt werden. Wird ein bestimmter Fehler von mehreren Testpersonen bemängelt, besteht akuter Handlungs- und Verbesserungsbedarf. Es ist davon auszugehen, dass bei einer Verwendung des Programms durch mehrere UserInnen, genau dieselben Probleme ebenfalls beanstandet würden (für Details siehe Kapitel 8.3).

Die Erkenntnisse der ProbandInnen wurden ebenfalls in der zweiten Phase und während der Lösungsversuche mittels der Thinking Aloud Methode gemessen. Dies ist die in dieser Magisterarbeit verwendete Methode, um Gedankenvorgänge von Testkandidaten und Probanden nachzuvollziehen und zu protokollieren. Thinking Aloud bedeutet so viel wie „lautes Denken“, dabei sollen die Testpersonen alle Gedankengänge, welche sie zum Zeitpunkt des Testvorgangs haben, egal ob positiv oder negativ, alle Ideen oder Gedankenfragmente laut auszusprechen (siehe Details hierfür im Kapitel 6). Alle getätigten Aussagen werden protokolliert, um Auswertungen zu ermöglichen. Aufgrund der großen Menge an Daten, die in besonders aktiven Phasen anfallen können, haben sich Videoaufnahmen bewährt.

Während die vier Aufgabenstellungen von den ProbandInnen bearbeitet wurden, diente der Autor dieser Magisterarbeit selbst als Moderator. Die beiden assistierenden Kollegen Herr Ulrich Fels und Herr Nikolaus Suchy hatten die Aufgabe, alle Vorgänge, Aktionen, Emotionen und auffällige Aussagen oder Aktionen zu notieren. Das gesamte Forschungsteam befand sich im selben Raum, wie die jeweiligen ProbandInnen, jedoch hinter deren Blickfeld. Das Ziel der zweiten Phase war es, die ProbandInnen die Aufgabenstellungen in einem ungezwungenen Umfeld selbstständig lösen zu lassen, während sie dabei ständig die Gedanken verbal laut und deutlich äußern. Im Idealfall musste der Moderator nach Ab-

schluss der Einführungsrede keine weiteren Hilfestellungen leisten oder Fragen beantworten.

Die Thinking Aloud Methode wurde wie zuvor erwähnt mit digitaler Bildschirmaufzeichnung (*Windows Media Encoder 9*) festgehalten. Diese Art der Beobachtung ist sehr zeitaufwendig und für die Interpretation und Kategorisierung von Benutzerhandlungen sind zusätzlich definierte Kriterien notwendig, siehe hierfür Kapitel 5.5 und 5.6 (Preece, et. al., 2002 zitiert in Rind, et. al., 2011). Die Dauer der Testdurchgänge belief sich auf eine bis zu zwei Stunden. Diese signifikante Schwankung entstand dadurch, dass einige ProbandInnen sich beträchtlich mehr Zeit bei Ihren Erkundungen nahmen, da die Aufgabenstellungen und das Interview keine zeitliche Begrenzung vorgaben. Aus den Aufzeichnungen wurden zwei Logfiles manuell und händisch vom Autor dieser Magisterarbeit selbst erstellt. Die erste Datei diente dazu alle Ereignisse, sowie Fehler und Unklarheiten während der Bedienung von TimeRider durch die jeweiligen UserInnen zu protokollieren. Diese Aufstellung soll in erster Linie der Untersuchung aller Usability relevanten Ereignisse dienen und die Fehler auf einen Blick erkenntlich machen. Die vollständigen Resultate der Usability Auswertung werden im Kapitel 8.5 vorgestellt und erläutert. Das zweite Logfile dokumentiert jegliche Aktivität der UserInnen, diese Daten könnten für eine Analyse nach Folgewahrscheinlichkeiten dienen, welche jedoch den Rahmen dieser Diplomarbeit sprengen würde.

In der dritten und letzten Phase wurde nach der Erfüllung aller Tasks ein kurzes Interview durchgeführt. Diese Befragungen dauerten in etwa jeweils 20 Minuten pro Testperson und es wurde ein Interviewleitfaden verwendet, damit sichergestellt werden konnte, dass die Befragung einheitlich durchgeführt wurden und alle wesentlichen Fragen abgedeckt wurden. Eine Beschreibung der Interview-Methode, der Vorgehensweise und der Interview-Resultate, wird ausführlich in der zweiten Magisterarbeit von Fels (2015) beschrieben. Die Untersuchungen und Interviews wurden von und mit jeder Testperson einzeln durchgeführt.

Nach den drei Phasen wurden die Ergebnisse der Studie, welche die Videodateien der Bildschirmaufzeichnungen darstellten, ausgewertet. Aus den Videos wurden vom Autor dieser Magisterarbeit, wie zuvor beschrieben zwei Logfiles erstellt. Die Videodateien, sowie die separaten Audioaufzeichnungen des Interviews dienten dann zur Utility-Analyse und zur eigentlichen Usability-Analyse. Für die Usability Fehleranalyse wurden die Kategorien von Forsell & Johansson (2010) als Grundlage herangezogen und für die Kategorisierung der Userinteraktivitäten wurden die Kategorien von und Yi, et. al. (2007) verwendet, welche es ermöglichen alle Ereignisse und Aktivitäten zu analysieren. Aus den Ergebnissen dieser Analysen konnten dann Rückschlüsse auf den inneren Interaktionsprozess während der Benutzung von TimeRider und auf Usability-Fehler gemacht werden, welche wiederum bei der Fehleranalyse des Programms hilfreich waren.

## 8.2 Beschreibung des Programms TimeRider

In diesem Kapitel wird das Programm TimeRider beschrieben und dessen dynamische Darstellung von visuellen Patientendaten einer Kohortenstudie, über die Zeit. Die Entwicklung von TimeRider wurde durch eine Kooperation der Technischen Universität Wien, der Donau-Universität Krems und mehreren niederösterreichischen Landesstellen für Gesundheit ermöglicht (Homepage des Projekts „VisuExplore“, 2012). Das Tool ist Bestandteil des „CVASt“ (*Centre for Visual Analytics Science and Technology*), welches vom Institut für Gestaltungs- und Wirkungsforschung und dem Institut für Software Technologie und interaktive Systeme an der Technischen Universität Wien ins Leben gerufen wurde.

TimeRider ist eine „Java WebStart“ Applikation<sup>16</sup> und verwendet bzw. erweitert die Visualisierungsmethode Animated Scatter Plot (siehe Kapitel 4.4) für die Analyse von Patientendaten einer Kohortenstudie, über die Zeit. Diese Masterarbeit welche TimeRider in der Programmversion 3.4 mit der Animated Scatter Plots Visualisierung auf Usability Fehler untersucht, diente als Grundlage für das wissenschaftliche Paper von Rind, et al. (2011). Darin werden die Informationen kurz und zusammengefasst vorgestellt, welche in den folgenden Kapiteln dieser Arbeit ausführlich beschrieben werden. Rind, et al. (2011) beschreiben das Forschungsfeld von Animationen in Visualisierungen als ein interessantes Forschungsfach, welches jedoch widersprüchliche Ergebnisse in vergangenen Studien aufweist (siehe Kapitel 4.2) (Rind, et al., 2011).

In TimeRider werden Patientendaten aus Kohortenstudien, über die Zeit dargestellt. Durch die Beobachtung des Verlaufs der Variablen über die Zeit, können Erkenntnisse gewonnen werden welche eine Analyse des Zusammenhangs und der Entwicklung der beiden Variablen über die Zeit ermöglicht. Die Variablen werden in einem kartesischen Koordinatensystem dargestellt. Die Schnittpunkte dieses Systems werden mittels Scatter Plot (Streudiagramm) dargestellt und sind der zentrale Bestandteil des Programms (siehe Abbildung 25). Auf den beiden Achsen wird jeweils nur eine Variable abgebildet. Die UserInnen können über ein „dropdown“ Menü, welches sich neben der jeweiligen Achse befindet, die gewünschte Variable auswählen (siehe Abbildung 26). Die Zuweisung der zwei gewünschten Variablen zu den Achsen, ermöglichen den UserInnen danach bis zu fünf weitere Variablen darstellungen zu wählen. Diese Optionen sind auf Größe, Farbe und Form beschränkt und verändern jeweils nur die Scatter Plot Datenpunkte welche der Auswahl entsprechen (siehe Abbildung 27).

---

<sup>16</sup> Weiterführende Materialien und das Programm TimeRider selbst (auszuführen via *Java WebStart*) sind zu finden unter: <http://www.cvast.tuwien.ac.at/projects/timerider>.

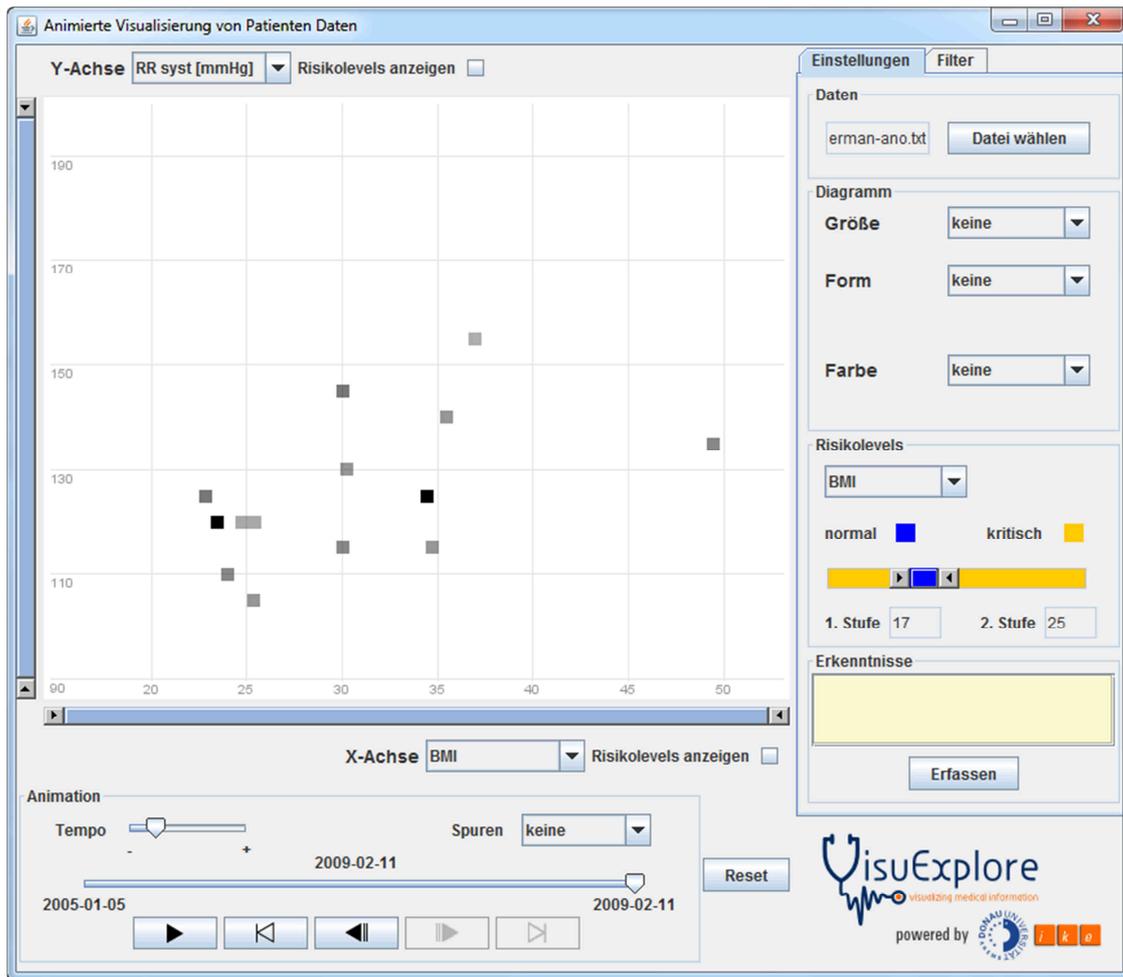


Abbildung 25: TimeRider Benutzerschnittstelle (User Interface/Frontend/GUI)<sup>17</sup>

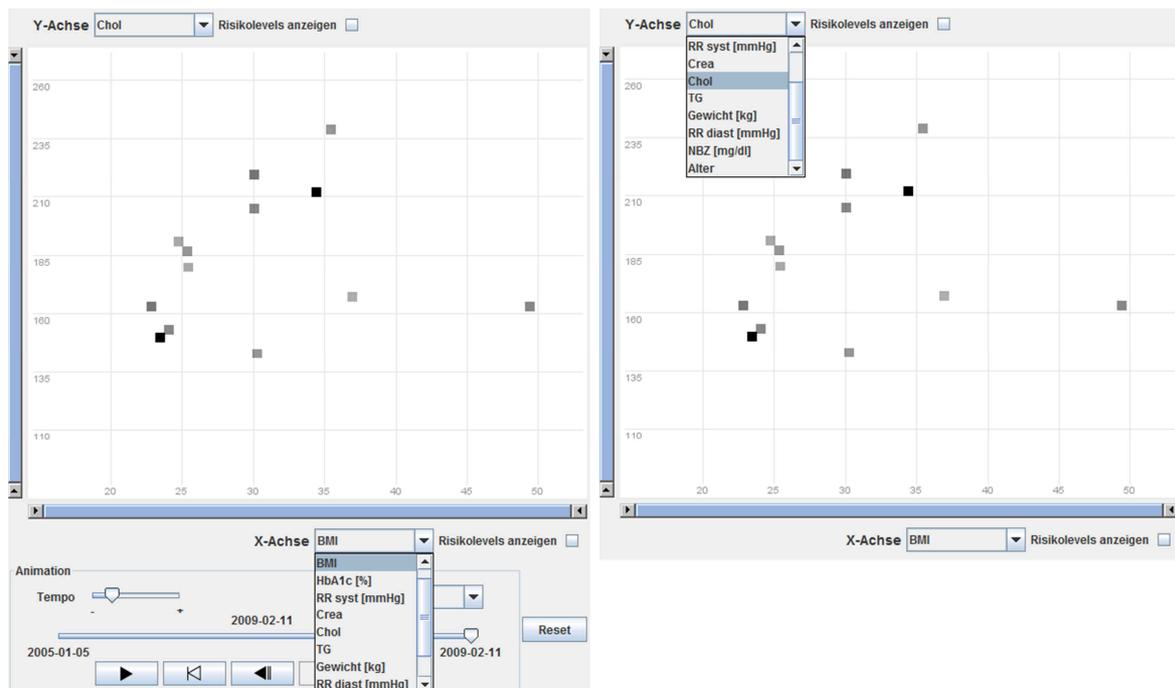


Abbildung 26: Achsen „dropdown“Menü in TimeRider<sup>18</sup>

<sup>17</sup> Der Screenshot zeigt die Darstellung von Patientendaten als Punkte in einem Streudiagramm (*scatter plot*).

<sup>18</sup> Das linke Bild zeigt das dropdown Menü der X-Achse und das rechte Bild zeigt das der Y-Achse.

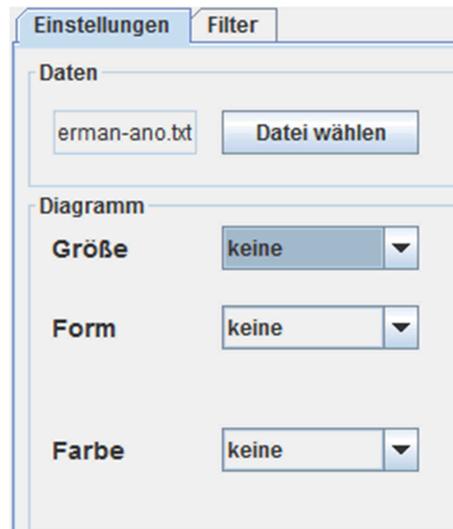


Abbildung 27: Variablendarstellung in TimeRider<sup>19</sup>

Die Animation der Daten wird verwendet um die Dynamik von Kohorten zu erforschen. Die Unregelmäßigkeit der gesammelten Daten wird durch eine einheitliche Zeiteinheit (z.B. ein Tag) vereinheitlicht, welche für die Darstellung der animierten Bilder diene. Bei jedem Bild wird ein Patient, sofern medizinische Daten über diesen an dem jeweiligen Datum vorhanden sind, als grafischer Punkt angezeigt. In jedem Bild zum jeweiligen Zeitpunkt wird eine Weg-Zeit-Linie zwischen vorhergehenden und nächsten Messungsdatenzeitpunkt gezogen.

Es wurden zwei visuelle Techniken in TimeRider verwendet, Traces und Transparenz um den Datenverschleiß und dem zeitlichen Kontext Rechnung zu tragen. Mit Transparenz ist der grafische Effekt gemeint, welcher bei der TimeRider Anzeige zu sehen ist. Dieser ist dadurch visualisiert, dass die Datenpunkte langsam über die Zeit ausblenden (ausgrauen=immer heller werden) wenn sie sich von der alten räumlichen Position (=Messzeitpunkt) wegbewegen. Sie sind erst dann vollkommen undurchsichtig, wenn in einem angezeigten Bild (=zu einem gewissen Zeitpunkt) auch Messungsdaten vorhanden sind. Dieser Effekt hebt Patienten mit aktuellen Daten hervor.

Im Trace-Modus wird eine durchgängige Linie gezogen, von der ersten räumlichen Position in der zu einem gewissen Zeitpunkt auch Messdaten vorhanden sind, bis zur nächsten Position an der eine Messung stattfand. Die Linie wird iterativ länger und zeigt immer alle vergangenen Positionen an und bleibt auch nach der letzten Position sichtbar. Am Ende der Animation und somit des gesamten Untersuchungszeitraumes, kann auf einen Blick der gesamte Verlauf der Messwerte über die Zeit und unter den gewählten Achsen-Werten betrachtet und analysiert werden. Dieses letzte statische Bild zeigt natürlich nur die Richtungsänderungen, sowie lokale Extremwerte an, sofern sie sich nicht gegenseitig überlagern.

<sup>19</sup> Die „dropdown“ Menüs (Größe, Form, Farbe), enthalten verschiedene auswählbare Variablen.

Animation wird dafür benötigt, um die Zeitwahl und zeitgleiche Darstellung dieser Entwicklungen zu verstehen. Die UserInnen können in der Zeit navigieren, indem sie über die Zeitachsen Funktionen (*play, stop, pause, fast forward, rewind*) den Ablauf steuern. Der Betrachter kann bequem über eine Play-Taste der Animation folgen und wie sich die Ereignisse über die Zeit entwickeln, oder eigenständig von Bild zu Bild und wieder zurück springen. Die Funktionstasten in TimeRider ermöglichen sowohl eine einfache Bedienung wie bei einem Media-Player oder die UserInnen verwenden den Zeitschieber (*timeslider*) manuell. Dieser kann mittels der Bewegung des Mauszeigers, indem der Schieberegler angeklickt wird und die Taste gleichzeitig gehalten und der Regler gezogen wird, die Abspiel-Geschwindigkeit regulieren und zwischen den Bildern hin und her wechseln (siehe Abbildung 28).

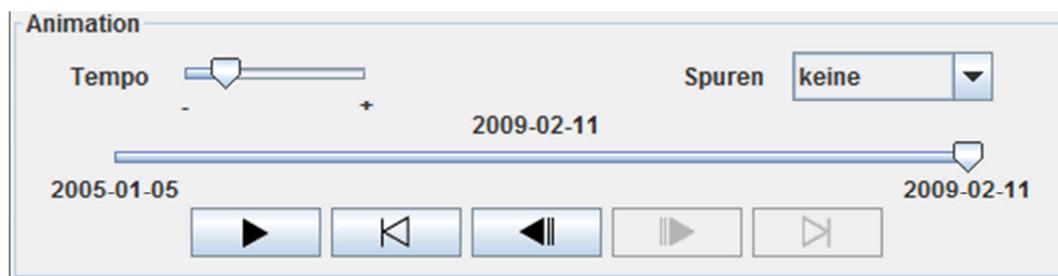


Abbildung 28: „timeslider“ Funktion in TimeRider<sup>20</sup>

Die meisten Variablen in der Humanmedizin haben wichtige Wertebereiche. Der Wert „HbA1c“ z.B. ist ein Glykohämoglobin und bedeutet, dass an den roten Blutfarbstoff (Hämoglobin) Glukose gebunden ist (Glykierung). Der Anteil des HbA1c am gesamten Bluthämoglobin (Hb) kann bei Untersuchungen des Patientenblutes festgestellt werden und wird meist als Prozentwert angegeben. Dieser gibt den HbA1c-Spiegel im Blut als Anteil am Gesamthämoglobin in Prozent an und hat in der Regel einen Normalwert von 4% bis 6%. Wenn dieser Wert überschritten wird, ist der Patient dem Risiko des Organversagens durch Diabetes ausgesetzt. Wenn dieser Grenzwert bereits in der Visualisierung hervorgehoben wird, kann eine schnellere Erkennung solcher Risikopatienten verbessert werden. In TimeRider wurde die Hervorhebung der Risikolevels durch eine hellblaue Hintergrundfärbung des Streudiagramms umgesetzt (siehe Abbildung 29). Die Risikoschwellen, welche in TimeRider hinterlegt sind, wurden von ExpertInnen festgelegt und entsprechen dem humanmedizinischen Standard. Die UserInnen können die Risikobereiche jeder Achse getrennt aktivieren. Die ÄrztInnen können nach der Aktivierung die Risikolevels jeweils adaptieren, um die Visualisierung nach Ihren Bedürfnissen zu optimieren.

<sup>20</sup> Die Animationsfunktion enthält einen Zeitschieber (*timeslider*), eine Geschwindigkeitssteuerung der Animation (*Tempo*), die Spurendarstellung (*Spuren*) und die Media Player Tasten. Die Play-Funktion wird nach Aktivierung automatisch zur Pause-Taste, rechts daneben ist die „spring zum Anfang“-Taste, in der Mitte die „Schritt zurück“-Taste, gefolgt von der „Schritt vorwärts“-Taste und abschließend die „spring zum Ende“-Taste.

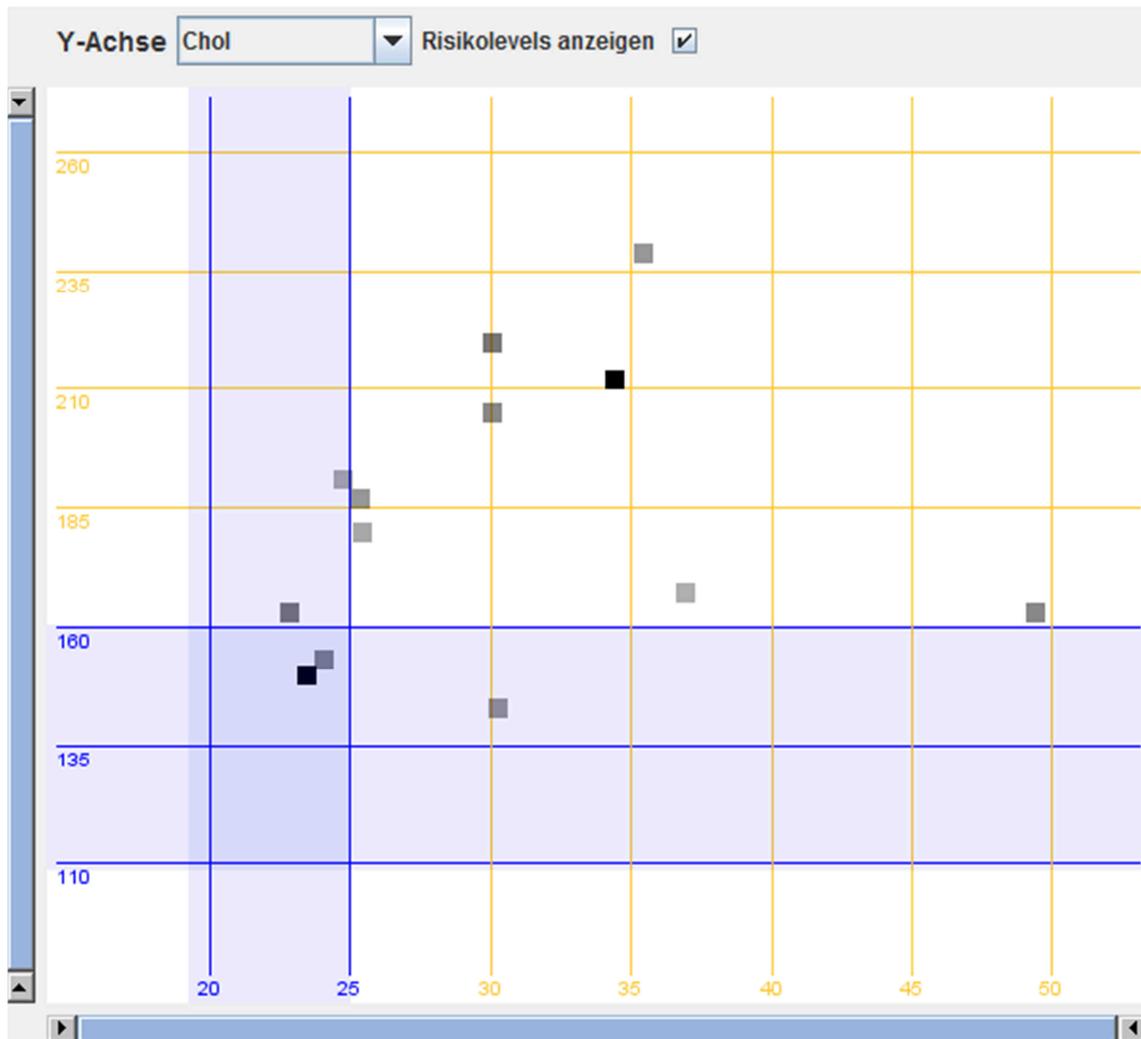


Abbildung 29: „Risikolevels“ Funktion in TimeRider<sup>21</sup>

TimeRider unterstützt übliche Interaktionen wie, verschieben eines Bildausschnitts (*panning*), zoomen, filtern und Detailansicht auf Abruf (*details on demand*, siehe Abbildung 30). Des Weiteren kann ein interaktiver Screenshot (Bildschirmfoto) des aktuellen Untersuchungsfortschritts angefertigt werden, dieser speichert alle Einstellungen für eine spätere Verwendung in eine Datei. Dieser Wiederherstellung des Zustands kann jedoch auch als Ausgangspunkt für eine andere Herangehensweise der Untersuchung verwendet werden oder mit Kollegen geteilt werden. Ebenfalls implementiert wurde die Funktion einen einfachen Screenshot zu machen und diesen als Bilddatei zu speichern. Eine *reset*-Funktion (Rücksetzfunktion) ergänzt das System, in dem das Programm in den Ursprungszustand zurückversetzt wird, indem es alle getätigten Einstellungen verwirft, falls ein Neuanfang gewünscht ist.

<sup>21</sup> Die blauen Bereiche markieren die Normwertebereiche und die gelben Bereiche sind die Risikobereiche. Die Datenpunkte bewegen sich je nach gemessenem Wert, während der Animation zwischen und innerhalb der beiden Bereiche. Es kann eine schnelle Übersicht gewonnen werden wie sich der Punkteschwarm verhält, indem das Verhalten über die Zeit beobachtet wird.

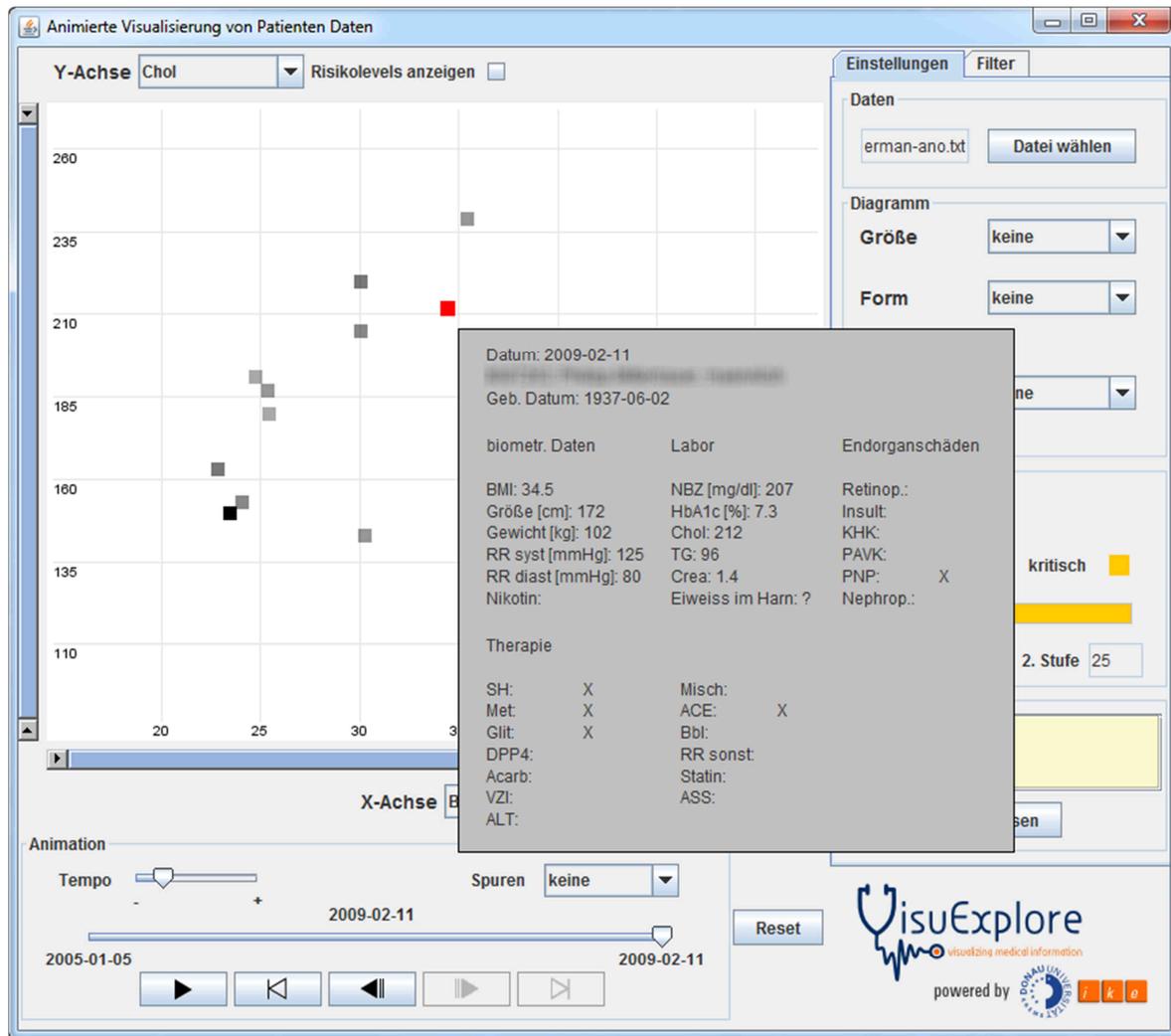


Abbildung 30: „details on demand“ Funktion, mittels „mouseover“ in TimeRider<sup>22</sup>

In dieser Arbeit wurden zwei verschiedene Kategorisierungssets verwendet, einerseits die zehn Kategorien von Forsell & Johansson (2010), welche zur Kategorisierung von Usability Problemen dienen und andererseits die sieben Kategorien nach Yi, et. al. (2007), welche einen Einblick in den Interaktionsprozess während der Benutzung von Informationsvisualisierungssystemen geben. Beide Sets haben das Ziel, die Untersuchungsergebnisse für die Fehleranalyse transparenter zu machen und Untersuchungen von verschiedenen Informationsvisualisierungssystemen vergleichbar zu machen.

Das Programm TimeRider wurde auch weiter entwickelt. In Neubauer (2012) wurde es um die *Small Multiple* Visualisierung (siehe Exkurs im Anhang C) sowie um die bestehende Möglichkeit zur Aktivierung der Traces (Linien-Darstellung zwischen den Scatter Plots bei aufeinanderfolgenden Untersuchungszeitpunkten) erweitert. Das Ziel dieser Arbeit war, das Verhältnis von zwei Variablen zueinander und deren Entwicklung über die Zeit zu analysieren und dabei zwischen einer dynamischen und zwei statischen Visualisierungen

<sup>22</sup> Diese Abbildung zeigt den „mouseover“ oder „hovering“ Effekt: Wenn der Benutzer den Mauszeiger über den Datenpunkt führt, „klappt“ ein Kontextmenü auf. Darin sind weitere Details des im Fokus stehenden Datenpunkts (=Patient) enthalten und können einfach überblickt werden.

empirische zu vergleichen. Die dynamische Visualisierung ist dabei die *Animated Scatter Plot* Darstellung (siehe Kapitel 4.4 für Details) und die statischen Visualisierungen sind einerseits die neu implementierte *Small Multiples* Darstellung (siehe Abbildung 31) und andererseits die der Traces (siehe Abbildung 32). Die beiden statischen Repräsentationen stellen die Daten in unterschiedlichen Zeitpunkte als kleine nebeneinander liegende Bilder dar (Neubauer, 2012).

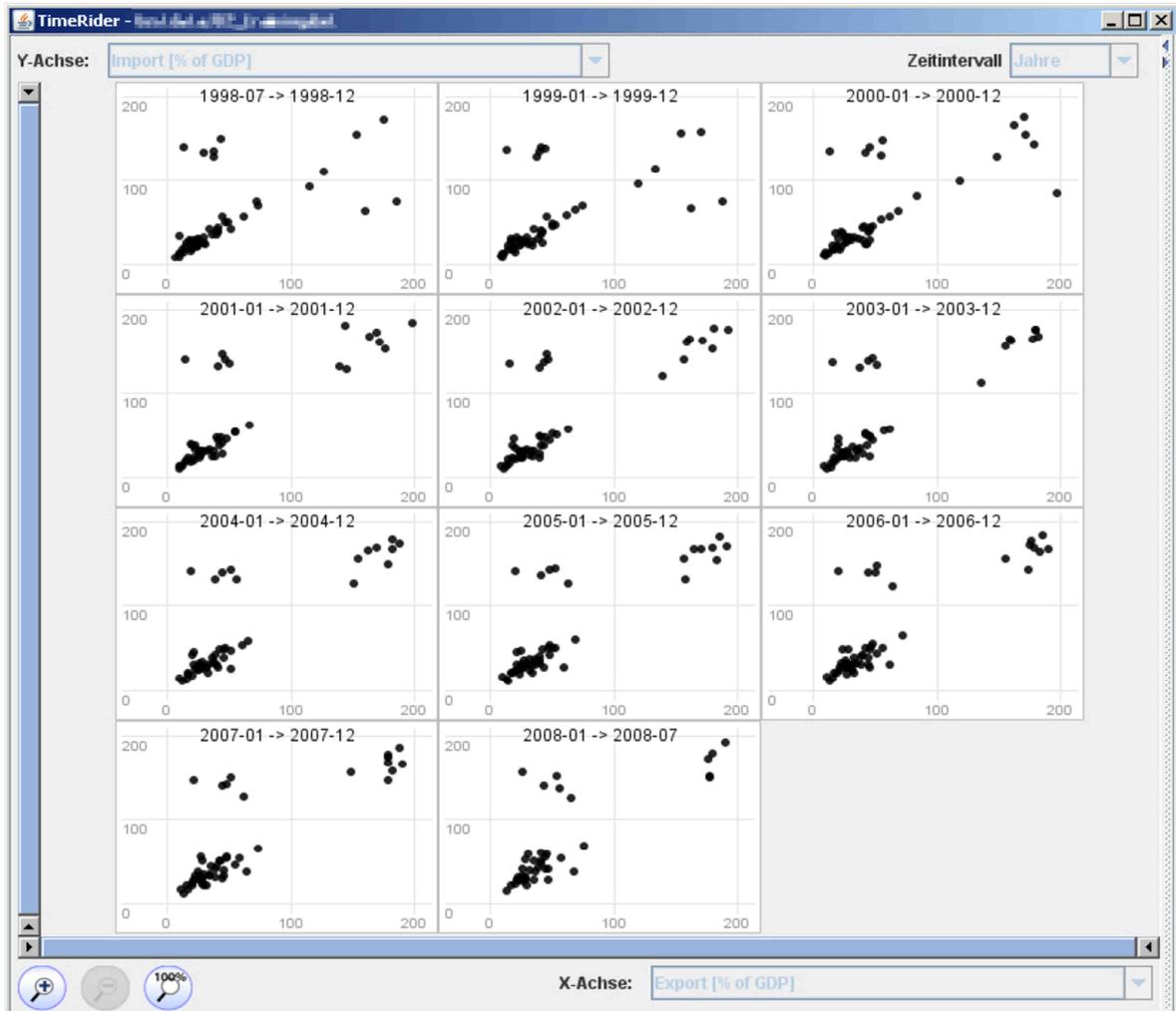


Abbildung 31: *Small Multiples* Visualisierung von TimeRider

Die Abbildung 31 zeigt eine *Small Multiples* Visualisierung von TimeRider. Es wird das prozentuale Verhältnis des Bruttoinlandsprodukts eines Landes über die Zeit dargestellt, in dem die Exporte (X-Achse) und Importe (Y-Achse) gegenübergestellt werden. Der Ablauf angefangen vom ersten Bild links bis zum letzten Bild unterste Reihe Mitte, stellt einen Zeitraum von zehn Jahren dar (vgl. Neubauer, 2012, p.64).

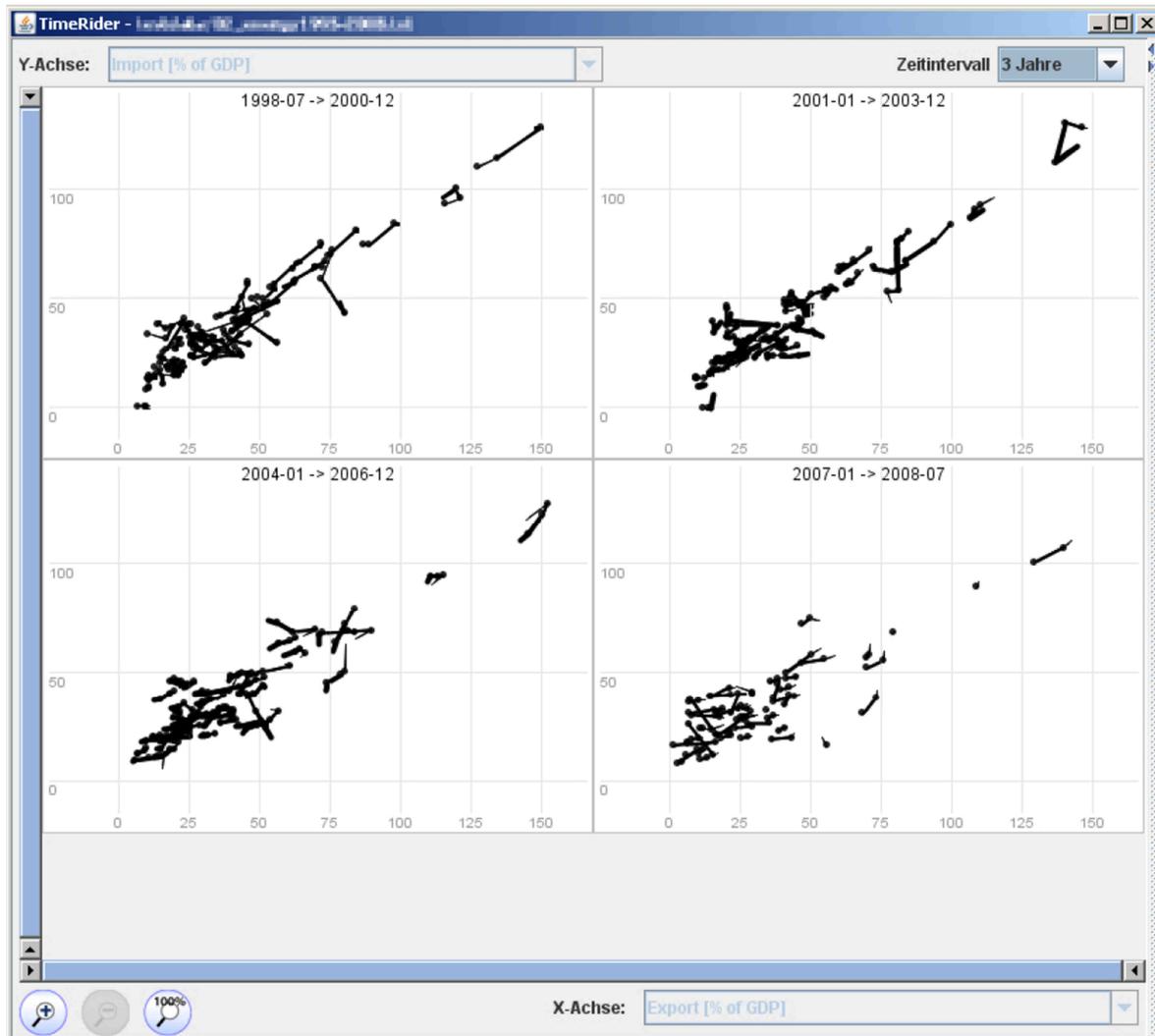


Abbildung 32: Small Multiples Visualisierung mit Traces von TimeRider

Die Abbildung 32 zeigt eine *Small Multiples* Visualisierung mit Traces von TimeRider. Die Bilder zeigen die Exporte (X-Achse) und Importe (Y-Achse) im prozentualen Verhältnis zum Bruttoinlandsprodukt über die Zeit. Angefangen vom ersten Bild links bis zum vierten und letzten Bild untere Reihe rechts, stellt einen Zeitraum von zehn Jahren dar (vgl. Neubauer, 2012, p.65).

### 8.3 ProbandInnen

In diesem Kapitel werden die ProbandInnen und einige Details beschreiben, soweit es die Anonymisierung zulässt. Das ProbandInnen-Kollektiv bestand aus 10 ÄrztInnen, davon vier Frauen und sechs Männer im Alter zwischen 26 und 35, mit einer einzelnen Person um die 50 Jahre alt. Diese Testpersonen waren vorab nicht in die Entwicklung von TimeRider miteinbezogen und sahen die Software zum ersten Mal. Vor Beginn des Tests wurden jeweils die ProbandInnen in die Funktionsweise von Thinking Aloud und auch in die Funktionsweise der Programmoberfläche eingewiesen. Besonderes wurde darauf hingewiesen,

dass das Programm TimeRider und seine Funktionsweise untersucht wird und nicht die Testperson selbst oder Ihre Fähigkeiten in Sachen Bedienung des User Interface. Der Autor dieser Masterarbeit und sein Testteam gestaltete die Einführung so, dass jeder Testperson die Programmoberfläche und die Methode Thinking Aloud mit einheitlichen Worten und Aktionen vorgeführt wurde.

Die ÄrztInnen stellten sich freiwillig und unentgeltlich zur Verfügung. Aufgrund des privaten Netzwerks des Autors konnten ausreichend ÄrztInnen motiviert werden, diese neuartige Software für den medizinischen Bereich zu testen. Diese ÄrztInnen waren in diversen Fachgebieten der Medizin tätig, konnten sich aber in das Themengebiet Diabetes schnell und gut einarbeiten. Die Dauer der Tests belief sich je nach Testperson auf ein bis zwei Stunden. Da die Aufgabenstellungen und das Interview keine zeitliche Begrenzung vorgeben, entstand diese signifikante Schwankung, da sich einige ProbandInnen mehr Zeit bei Ihren Erkundungen nahmen als die Anderen.

Nielsen weist auf der Homepage „NN/g“ (2000) darauf hin, dass Usability Untersuchungen nicht kostenintensiv und Komplex sein müssen. Er geht sogar so weit, dass er aufwändige Usability Tests als „Verschwendung von Ressourcen“ (Homepage der „NN/g“, 2000) bezeichnet. Die besten Resultate erzielt man laut Nielsen wenn man nicht mehr als fünf BenutzerInnen testet und dafür so viele Tests mit geringem Aufwand, wie leistbar durchgeführt (Homepage der „NN/g“, 2000). Nielsen & Landauer (1993) beschrieben in ihrer Arbeit das Verhältnis zwischen gefundenen Usability Problemen und Test-UserInnen, das Ergebnis ist in der nachfolgenden Abbildung 33 dargestellt:



Abbildung 33: Verhältnis zwischen gefundenen Usability Problemen und Test-UserInnen<sup>23</sup>

Eindeutig zu sehen ist, dass wenn keine BenutzerInnen getestet werden, dann gibt es auch keine gefundenen Usability Probleme. Nur durch bereits ein(e) UserIn steigt die Anzahl

<sup>23</sup> Der Screenshot zeigt die steigende Anzahl der gefundenen Usability Probleme im Verhältnis der Anzahl von Test-UserInnen (Homepage der „NN/g“, 2000).

der gefunden Usability Fehler auf ein Drittel von allen möglichen Usability Problemen im Programmdesign. Beim Hinzuziehen einer zweiten Testperson werden weitere Usability Probleme aufgedeckt, welche von nur einer Testperson übersehen wurden, auch wenn die Steigerungsrate bereits nicht mehr so hoch ist. Ab drei Test-UserInnen ist der Lernerfolg immer geringer und es ist kosteneffizienter maximal fünf UserInnen kleine Tests durchführen zu lassen und diese Ergebnisse gleich in „Redesigns“ der Software einzuarbeiten und wieder fünf UserInnen testen lassen, usw. Dabei muss jedoch bedacht werden, dass die Behebung von Usability Fehlern, neue erzeugen kann, welche beim neuerlichen Testen dann zum ersten Mal erkannt werden. Laut Nielsen verschwendet man ab fünf UserInnen nur Zeit und Geld, da sich die gravierendsten Probleme wiederholen und man bereits 85% aller Usability Fehler entdeckt hat (Homepage der „NN/g“, 2000).

Die Grafik zeigt aber auch, dass zumindest 15 UserInnen benötigt werden um alle Usability Probleme des aktuellen Programmdesigns zu finden. Ab dem zweiten Test können dann auch tiefgehende Usability Funktionen des Programms untersucht werden, wie fundamentale Strukturen, Informationsarchitektur, Aufgabenablauf und Bedürfnisse der UserInnen, da nach ein, zwei kleinen Tests die meisten Usability Fehler entdeckt und behoben sind. Der dritte Test ist dann eine Art letzter Schliff, der überprüfen soll ob die Überarbeitungen fehlerfrei funktionieren und ob weitere tiefgreifende Usability Probleme, wie zuvor beschrieben, entdeckt werden (Homepage der „NN/g“, 2000).

Da in dieser Magisterarbeit 10 ProbandInnen getestet wurden, sei an dieser Stelle darauf verwiesen, dass unser Anspruch erstens nicht ökonomisch orientiert war, somit war unser Anspruch die größte mögliche Anzahl an Usability Problemen zu finden. Da es schwierig war 15 Professionals zu motivieren um sich für einen mehrstündigen Usability-Test Zeit zu nehmen wurde der Kompromiss gefunden, dass 10 Testpersonen ausreichend sind, da laut Nielsen diese zwischen 96% und 98% aller Usability Probleme finden werden.

## 8.4 Fragestellung

Dieses Kapitel erläutert die Fragestellung dieser Arbeit und die daraus resultierenden vier Hypothesen dieser Magisterarbeit. Die zentrale Fragestellung ist ob nach einer kurzen Einarbeitungszeit mittels verschiedener Aufgabenstellungen ein eigenständiges Arbeiten mit dem Programm möglich ist. Dies ist von didaktischem Interesse, da dieses Programm auch in der Ausbildung von ÄrztInnen und medizinischen Personal eingesetzt werden soll. Das rasche Erlernen ist auch deshalb wichtig, weil ÄrztInnen wenig Zeit haben um sich in komplexe Programme einzuarbeiten und daher ein möglichst einfaches Programm gewünscht wird (Pohl, et al., 2011).

Die Ausbildung und das Training von Mediziner\*innen sind sehr zeitaufwändig und teuer. Deshalb wird in der Humanmedizin von der theoretischen Grundausbildung bis hin zum Erwerb handwerklicher Fähigkeiten verstärkt auf Simulation mit interaktiver 3D-Computergrafik gesetzt. Ein Überblick über Virtual Reality Techniken in der humanmedizinischen Aus- und Weiterbildung ist in Cakmak, et al., (2000 zitiert in Nischwitz, et al., 2011) zu finden (Nischwitz, et al., 2011). Der Lernerfolg wurde mittels der letzten von vier Aufgabenstellungen zur Untersuchung des Programms und in den nach der Thinking Aloud Studie durchgeführten Interviews überprüft.

Für die Evaluierung der Usability von TimeRider konnten, wie bereits im Kapitel 8.3 zuvor erwähnt, 10 Ärzt\*innen (*Professionals*) als Tester gewonnen werden. Im Gegensatz zu den meisten anderen Studien, welche nur Medizinstudent\*innen als Testpersonen heranzogen (vgl. Pohl, et al., 2012; Martins, et al., 2008). Die Literaturrecherche für diese Masterarbeit, die Planung für die Tasks und die Einarbeitung in TimeRider ergaben die nachfolgenden vier Hypothesen:

1. **Hypothese 1:** Eigenständiges Arbeiten mit dem Programm ist bereits nach einer kurzen Einschulung und Einarbeitungszeit mit verschiedenen Aufgabenstellungen möglich.

Die erste Hypothese überprüft die zentrale Annahme dieser Masterarbeit, dass ein Lerneffekt bei TimeRider nach kurzer Einschulung und Einarbeitung erzielt werden und das Programm dann eigenständig verwendet werden kann. Wie am Anfang dieses Kapitels beschrieben, ist dies auch von didaktischem Interesse, da dieses Programm auch in der Ausbildung von Ärzt\*innen und medizinischen Personal eingesetzt werden soll. Durch eine kürzere Einschulungs- und Einarbeitungsphase, könnte dadurch im klinischen Alltag eine breite Akzeptanz erreicht werden, da das Programm schnell und intuitiv zu genutzt werden kann.

2. **Hypothese 2:** Die Animation der zeitorientierten Daten in TimeRider hilft den Ärzt\*innen einen Trend zu erkennen.

Die zweite Hypothese überprüft die Annahme, dass die Animation der zeitorientierten Daten für die Ärzt\*innen eine Unterstützung in der Patientendatenanalyse darstellt. Wichtig war daher ob ein Trend zu erkennen ist wenn die Punktwolke des Streudiagramms sich über die Zeit verändert. Diese intuitive Trenderkennung ist ein Vorteil in der medizinischen Ausbildung und im klinischen Alltag, da durch die Verwendung von TimeRider eine Verringerung des zeitlichen Aufwands bei Patientendatenanalysen zu erwarten ist. Dies wiederum ist, wie schon erwähnt, erstrebenswert, da die Ausbildung und das Training von Mediziner\*innen sehr zeitaufwändig und teuer sind.

3. **Hypothese 3:** Das Mapping der Daten (Scatter Plot, Farbe, Traces, usw.), wie sie in TimeRider implementiert wurde, unterstützt die ÄrztInnen bei der Lösung der Aufgabenstellungen.

Die dritte Hypothese soll die Annahme überprüfen, dass das gewählte Mapping der Daten in TimeRider, die Testpersonen bei der Bearbeitung der gestellten Aufgaben unterstützt. Sollten die ÄrztInnen, welche eine spezielle Klientel darstellen, die fundamentale Darstellung nicht intuitiv akzeptieren, dann müsste eine vollständige Überarbeitung des Programms erfolgen. ÄrztInnen werden schon im Studium auf bestimmte Diagnose- und Analyseverfahren, sowie eigene Farbkodierungen, Falschfarbenbilder, usw., trainiert und deshalb ist eine Akzeptanz des Daten-Mapping bei neuen medizinischen Diagnose- oder Schulungsprogrammen essentiell.

4. **Hypothese 4:** Die ÄrztInnen entdecken, wie von Nielsen (Homepage der „NN/g“, 2000) beschrieben, Usability und Utility Fehler.

Die vierte Hypothese überprüft die Annahme inspiriert von Nielsen, dass die 10 Testpersonen nahezu alle Usability und Utility Fehler finden. Diese Fehler und Schwierigkeiten sollten in ein „Redesign“, wie von Nielsen (siehe Kapitel 5.3) empfohlen, einfließen, welches dann wieder mittels Usability Untersuchung getestet werden sollte. Nielsen empfiehlt drei Durchgänge, da danach anzunehmen ist, dass der Großteil der ursprünglichen Fehler und die des „Redesign“ entdeckt wurden und behoben werden konnten. Diese zusätzlichen Forschungen würden den Rahmen dieser Diplomarbeit sprengen und obliegen daher zukünftigen Untersuchungen.

## 8.5 Resultate

In diesem vorletzten Kapitel werden nun alle beobachteten und berechneten Ergebnisse des Usability Tests präsentiert. Es folgen nun alle daraus gemessenen und relevanten Fakten der vier Hypothesen und danach eine Auflistung aller entdeckten Usability und Utility Fehler anhand der Kategorien von Forsell & Johansson (2010) und alle kategorisierten Interaktionen nach Y, et al. (2007). Danach folgt eine Zusammenfassung der Fakten und Erfahrungen der angewandten Thinking Aloud Methode.

### 8.5.1 Übersicht

Die Zielsetzung dieser Magisterarbeit war es, den Lerneffekt bei ÄrztInnen durch das Programm TimeRider zu untersuchen, indem nach einer kurzen Einführung und einer Einarbeitungsphase die ProbandInnen selbständig mit dem Programm arbeiten mussten. Dazu wurde ein Usability Test mittels der Thinking Aloud Methode durchgeführt, welcher ins-

gesamt 7Stunden 41Minuten Videomaterial erzeugte. Dieses Datenmaterial wurde vollständig vom Autor dieser Magisterarbeit manuell ausgewertet und ein Logfile angefertigt, welches jede Aktion der ProbandInnen und die dadurch entdeckten Usability Fehler dokumentierte. Es wurde angenommen, dass die ProbandInnen während der Lösung von vier Aufgabenstellungen einem Usability Test mittels Thinking Aloud Methode unterzogen werden können und durch die gewonnen Daten ein Lerneffekt nachgewiesen werden konnte. In den Usability Tests wurden Fehler im Design des Programms hinsichtlich „Usability“ (= Benutzerfreundlichkeit eines Programmes) und „Utility“ (= Nutzen eines Programmes) untersucht und beschrieben. Diese Erkenntnisse sollen in zukünftige Aktualisierungen (*Updates*) des Programmes einfließen. Es folgen nun die Ergebnisse ob die vier Hypothesen bestätigt oder widerlegt werden konnten.

## 8.5.2 Hypothesen

### Hypothese 1

Alle ProbandInnen (10 von 10) hatten anfängliche Schwierigkeiten mit der anfänglichen Bedienung des Programms, da ihnen dieses bis zuvor unbekannt war. Es konnte jedoch beobachtet werden, dass alle ProbandInnen (10 von 10) nach drei Aufgabenstellungen die letzte und vierte Aufgabe TASK IV lösen konnten (siehe Anhang A.4), wobei diese keine Vorgaben enthielt und nur zum Experimentieren animierte (siehe Tabelle 1).

Tabelle 1: Auswertung der vierten Aufgabe (Lerneffekt).

TASK IV (Hypothese 1 - Lerneffekt)				
Testperson	TASK IV gelöst	Lösungszeit	TASK Parameter	Neue Parameter
ProbandInn 0	X	14m 19sec		X
ProbandInn 1	X	8m 59sec	X	
ProbandInn 2	X	7m 18sec	X	
ProbandInn 3	X	9m 57sec		X
ProbandInn 4	X	12m 04sec		X
ProbandInn 5	X	5m 54sec		X
ProbandInn 6	X	7m 38sec		X
ProbandInn 7	X	7m 32sec		X
ProbandInn 8	X	10m 28sec		X
ProbandInn 9	X	11m 03sec		X

Die Auswertung von TASK IV ist in der abgebildeten Tabelle 1 zusammengefasst und nach vier Kategorien gefiltert. Man sieht auf einen Blick, dass alle ProbandInnen die Aufgabenstellung lösen konnten und dafür bis zu 14 Minuten und 19 Sekunden aufwendeten. Des Weiteren wurde beobachtet das 2 von 10 ProbandInnen, die vierte Aufgabe dazu nutzen, gezielt einige Vorgaben der Parameter aus den vorhergehenden Tasks nochmals zu verwenden. Die anderen 8 von 10 ProbandInnen verwendeten neue Parameter, welche sie interessant fanden auf den kartesischen Achsen in Relation zu setzten, bzw. sie verwendete-

ten jene Parameter, die für eine Analyse von Diabetes Mellitus Patientendaten, neue Erkenntnisse beinhalten könnten. Eine Testperson erklärte sogar, dass durch ein vermehrtes Arbeiten mit dem Programm, eine bessere Orientierung in TimeRider erlangt werde.

In der ersten Hypothese wurde die Annahme beschrieben: „Eigenständiges Arbeiten mit dem Programm ist bereits nach einer kurzen Einschulung und Einarbeitungszeit mit verschiedenen Aufgabenstellungen möglich“. Alle ProbandInnen (10 von 10) hatten anfängliche Schwierigkeiten mit der Bedienung des Programms. Es konnte jedoch beobachtet werden, dass trotzdem alle ProbandInnen (10 von 10) nach drei Aufgabenstellungen in der vierten Aufgabenstellung sich intuitiv zu Recht fanden, obwohl diese letzte und vierte Aufgabe TASK IV (siehe Anhang A.4) keine Vorgaben enthielt und freies experimentieren forderte. Die ProbandInnen nutzten diese letzte „freie“ Aufgabenstellung um Zusammenhänge von Daten aus humanmedizinischer Sicht zu analysieren und verwendeten dafür selbständig und mehrheitlich neue Parameter die von Interesse waren. Somit wurde bewiesen, dass das Programm TimeRider sowohl für Schulungen und Ausbildungen, als auch zu Arbeit in Kliniken geeignet ist, da es schnell und intuitiv erlernbar ist. Die Annahme, dass sich durch Abarbeitung von Aufgaben, welche die Verwendung des Programms inkludierten, ein Lerneffekt einstellte, wurde durch die Ergebnisse dieser Beobachtung bewiesen und bestätigt somit die Hypothese 1.

## **Hypothese 2**

Für die Überprüfung der zweiten Hypothese wurde in TASK III (siehe Anhang A.4) beobachtet, dass 9 von 10 ProbandInnen den Zeitschieber verwendeten und nur eine Testperson die Media Player Funktionstasten, bevor diese ebenfalls den Zeitschieber verwendete. Mit Beginn der Videoaufnahme des Tests interagierten 7 von 10 ProbandInnen mit dem Zeitschieber innerhalb von 50 Sekunden, die kürzeste Zeit bis mit dem Zeitschieber interagiert wurde war 23 Sekunden, die längsten Zeit war 16 Minuten 11 Sekunden (siehe Tabelle 2). Nach diesen ersten Interaktionen mit dem Zeitschieber der Animationsfunktion wurde insgesamt von allen ProbandInnen 598-mal die Zeitfunktionen verwendet, wobei 44 Fehler und Ereignisse dokumentiert wurden (siehe weiter unten Auswertungen der Kategorie nach Yi, et al., 2007). Alle ProbandInnen (10 von 10) lösten TASK III, indem sie den geforderten Gruppentrend erkannten.

Die zweite Hypothese beschreibt die Annahme: „Die Animation der zeitorientierten Daten in TimeRider hilft den ÄrztInnen einen Trend zu erkennen“. Interessanterweise war zu beobachten, dass 9 von 10 ProbandInnen intuitiv den Zeitschieber verwendeten und nur eine Testperson die Media Player Funktionstasten verwendete, bevor sie ebenfalls den Zeitschieber verwendete. Nach diesen ersten Interaktionen mit dem Zeitschieber der Animationsfunktion wurde insgesamt von allen ProbandInnen insgesamt ca. 600-mal die Zeitfunktionen zur Lösung der Aufgaben verwendet. Diese häufige Verwendung der Zeitfunk-

tionen resultiert aus der Analyse von Gruppentrends mittels Animation der Scatter Plots. Daher ist es nicht verwunderlich, dass alle ProbandInnen (10 von 10) in TASK III (siehe Anhang A.4) den geforderten Gruppentrend erkannten und die Aufgabe lösen konnten. Während dieser einen Aufgabe konnte beobachtet werden, dass von der ersten Aufgabenstellung an, die Erkennung von Trends mittels Animation durch die ProbandInnen intuitiv und selbständig durchgeführt wurde. Durch diese eindeutigen Beobachtungen ist Hypothese 2 bestätigt worden.

Tabelle 2: Auswertung der dritten Aufgabe (Animation).

TASK III (Hypothese 2 - Animation)				
Testperson	TASK III gelöst	1. Zugriff	Zeitschieber	Funktionstasten
ProbandInn 0	X	16m 11sec	X	
ProbandInn 1	X	31sec	X	
ProbandInn 2	X	26sec		X
ProbandInn 3	X	1m 18sec	X	
ProbandInn 4	X	29sec	X	
ProbandInn 5	X	23sec	X	
ProbandInn 6	X	29sec	X	
ProbandInn 7	X	43sec	X	
ProbandInn 8	X	48sec	X	
ProbandInn 9	X	1m 19sec	X	

### Hypothese 3

Diese Annahme wurde in der zweiten Aufgabenstellung TASK II (siehe Anhang A.4) überprüft. Es wurde beobachtet, dass alle ProbandInnen (10 von 10) die zweite Aufgabe lösen konnten. Die Veränderung der Form der Scatter Plots wurde von 2 von 10 ProbandInnen durchgeführt und die Veränderung der Farbe von den anderen 8 von 10 ProbandInnen. Alle ProbandInnen (10 von 10) entdeckten innerhalb von 30 Sekunden die Funktionen für die Änderung von Farbe oder Form im Menü Einstellungen/Diagramm (siehe Tabelle 3). Innerhalb von 15 Sekunden änderten 8 von 10 ProbandInnen das Mapping der Scatter Plots im jeweiligen „dropdown“ Menü für Farbe, wie in der zweiten Aufgabenstellung ohne weitere Anleitung gefordert. Dabei bemängelten 3 von 10 ProbandInnen jedoch den Inhalt des „dropdown“ Menüs, da in diesem ohne erkennbare Sortierungslogik die Parameter angezeigt werden. Bei 2 von 10 ProbandInnen wurden erweiterte Menüoptionen für eine manuelle Auswahl der Farbe z.B. mittels einer Farbpaletten-Option erwartet.

Die Annahme der dritten Hypothese war: „Das Mapping der Daten (Streudiagramm, Farbe, Spuren, usw.), wie sie in TimeRider implementiert wurde, unterstützt die ÄrztInnen bei der Lösung der Aufgabenstellungen“. Diese Annahme wurde in der zweiten Aufgabenstellung TASK II (siehe Anhang A.4) überprüft. Es wurde beobachtet, dass alle ProbandInnen (10 von 10) die zweite Aufgabe lösen konnten. Am häufigsten wurde die Veränderung der Farbe der Scatter Plots gewählt (8 von 10 ProbandInnen), die restlichen Testpersonen ver-

änderten die Form der angezeigten Datenpunkte. Alle Test UserInnen (10 von 10) entdeckten innerhalb kürzester Zeit von 30 Sekunden die Funktionen für die Änderung von Farbe oder Form im Menü Einstellungen/Diagramm. Die meisten ProbandInnen (8 von 10) änderten innerhalb von 15 Sekunden intuitiv das Mapping der Scatter Plots, im dafür vorgesehenen Menü in TimeRider. Hypothese 3 wurde durch diese dokumentierten Ergebnisse bestätigt.

Tabelle 3: Auswertung der zweiten Aufgabe (mapping).

TASK II (Hypothese 3 - Mapping)				
Testperson	TASK II gelöst	1. Zugriff	Form	Farbe
ProbandInn 0	X	12sec		X
ProbandInn 1	X	14sec	X	
ProbandInn 2	X	9sec		X
ProbandInn 3	X	10sec	X	
ProbandInn 4	X	33sec		X
ProbandInn 5	X	13sec		X
ProbandInn 6	X	29sec		X
ProbandInn 7	X	14sec		X
ProbandInn 8	X	11sec		X
ProbandInn 9	X	5sec		X

#### Hypothese 4

Während des gesamten Usability Test wurden insgesamt 1.067 nach Forsell & Johansson (2010) kategorisierte Fehler gemessen und 3.700 Interaktionen nach Yi, et al. (2007) beobachtet. In Tabelle 4 ist die Zusammenfassung aller Beobachtungen von allen ProbandInnen dargestellt und nach der jeweiligen Kategorisierung-Methode aufgeteilt.

Tabelle 4: Eine Übersicht über alle kategorisierten Ereignisse von allen ProbandInnen.

Kategorien nach Yi, et al. (2007)		Kategorien nach Forsell & Johansson (2010)	
Alle ProbandInnen		Alle ProbandInnen	
Yi, et al.	Anzahl	Fors.&Joha.	Anzahl
abstract/ elaborate	1201	B3	197
Variablen	919	B5	175
Zeitachse	598	E1	172
filter	277	B9	116
encode	276	E7	105
reconfigure	244	D10	91
connect	135	E16	90
explore	49	E11	43
select	-	B7	42
		C6	36

Die vierte Hypothese deckte die Annahme ab: „Die ÄrztInnen entdecken, wie von Nielsen (Homepage der „NN/g“, 2000) beschrieben, Usability und Utility Fehler“. Die Auswertung des gesamten Usability Tests anhand der Kategorien von Forsell & Johansson (2010) erga-

ben insgesamt 187 Usability Probleme. Des Weiteren wurden 76 Programmfehler und 7 Systemabstürze gezählt, welche dringend behoben werden müssen. Die Hypothese 4 wurde somit eindeutig bestätigt.

### Usability Fehler:

In dieser Magisterarbeit wurden zwei Kategorie-Sets als Grundlage für die Interpretation und Untersuchung von Benutzeraktivitäten in TimeRider verwendet (siehe Kapitel 5.5 und 5.6), obwohl diese Rahmenbedingungen ursprünglich für heuristische Auswertungen erstellt wurden. Dadurch war es möglich eigenständige Ereignisse und Aktivitäten zu kategorisieren, wie z.B. wenn die UserInnen die Animation mittels direkter Manipulation der Zeitachse vor- und zurücklaufen ließen oder die Achsenparameter während eines Task plötzlich änderten, um einen anderen Blickwinkel zu bekommen.

Es stellte sich jedoch als schwieriger als gedacht heraus, da manche Usability Probleme nicht ganz einfach zu kategorisieren waren, da sich einige Fehler nicht eindeutig mit genau nur einer Kategorie beschreiben ließen. Damit mehrfach Kategorisierungen pro Fehler verhindert werden, wurde innerhalb des Forschungsteam (Herr Nikolaus Suchy, Herr Ulrich Fels BSc und der Autor dieser Arbeit selbst) versucht einen Konsens pro Fehler zu finden. Durch diese Vorgehensweise wurden insgesamt 187 Fehler nach Forsell & Johansson (2010) identifiziert, welche in Tabelle 5 von allen ProbandInnen dargestellt wird.

*Tabelle 5: Die überarbeitete Übersicht über alle kategorisierten Fehler nach Forsell & Johansson (2010) von allen ProbandInnen*

<b>Kategorien nach Forsell &amp; Johansson (2010)</b>		
<b>Alle ProbandInnen</b>		
<b>Forsell &amp; Johansson</b>	<b>Codierung</b>	<b>Anzahl</b>
<b>spatial organization</b>	<b>B3</b>	<b>84</b>
<b>consistency</b>	<b>E16</b>	<b>58</b>
<b>prompting</b>	<b>E1</b>	<b>29</b>
<b>flexibility</b>	<b>E11</b>	<b>5</b>
<b>recognition rather than recall</b>	<b>C6</b>	<b>4</b>
<b>minimal actions</b>	<b>E7</b>	<b>2</b>
<b>orientation and help</b>	<b>B7</b>	<b>2</b>
<b>information coding</b>	<b>B5</b>	<b>1</b>
<b>remove the extraneous</b>	<b>D10</b>	<b>1</b>
<b>data set reduction</b>	<b>B9</b>	<b>1</b>

Des Weiteren wurde von einer der ProbandInnen angemerkt, dass das „mouseover“-Kontextmenü aller Scatter Plots für den klinischen Einsatz überarbeitet werden muss, hinsichtlich welche Daten angezeigt werden und deren Nomenklatur.

### Programmfehler (*bugs*)/ -abstürze:

Es wurden insgesamt bei allen ProbandInnen 76 Programmfehler während der Usability Tests dokumentiert. Beim Test einer der ProbandInnen alleine wurden insgesamt 18 Programmfehler gezählt und bei einer der Anderen ProbandInnen wurden wiederum nur 2 Programmfehler dokumentiert. Die Tabelle 6 zeigt die restliche Anzahl von dokumentierten Programmfehlern und bei welchen ProbandInnen diese jeweils aufgetreten sind.

Tabelle 6: Eine Übersicht über aller Programmfehler/ Abstürze welche je nach Testperson auftraten

BUGs	
Testperson	Anzahl
Probandinn0	18
Probandinn7	12
Probandinn8	10
Probandinn1	7
Probandinn2	7
Probandinn3	7
Probandinn5	5
Probandinn6	5
Probandinn9	3
Probandinn4	2

Während der Usability Tests wurden auch insgesamt sieben Programmabstürze gezählt. Es wurden drei Programmabstürze bei „ProbandIn 0“, zwei Programmabstürze bei „ProbandIn 1“ und ein Programmabsturz bei „ProbandIn 2“ gezählt, bei den restlichen ProbandInnen kam es zu keinen dokumentierten Abstürzen (siehe in Tabelle 7, die Auflistung aller dokumentierten Abstürze).

Tabelle 7: Eine Übersicht über aller Programmabstürze welche je nach Testperson auftraten

Programmabsturz	
Testperson	Anzahl
Probandinn0	3
Probandinn1	2
Probandinn2	1
Probandinn7	1
Probandinn3	0
Probandinn4	0
Probandinn5	0
Probandinn6	0
Probandinn8	0
Probandinn9	0

### 8.5.3 Thinking Aloud (Usability aus Sicht der ProbandInnen Aktivitäten)

Das Thinking Aloud Protokoll nach Boren & Ramey (2000) konnte durch das Forscherteam intuitiv angewendet werden. Auch die vielen Tipps der beiden Autoren, wie die Bestätigungszeichen (*acknowledgment tokens*) oder wie man vorgeht, wenn eine der ProbandInnen gedanklich stecken geblieben ist, waren sehr hilfreich. Nach der ersten Durchsicht der Arbeit von Boren & Ramey, wird nach Meinung des Autors dieser Magisterarbeit aber ein anderes Bild des Ablaufs des Thinking Aloud Protokolls vermittelt. Es wird zwar darauf verwiesen, dass vor allem Usability Tester die Anweisungen sehr großzügig auslegen würden, da sie die Aufgabe haben Fehler zu finden, aber nichts kann jemanden darauf vorbereiten, wie schwierig es ist als menschliches Wesen die natürliche Sprachkommunikation zu unterdrücken und nur in einer sehr rudimentären Art und Weise dem Gegenüber (Testperson) zu antworten bzw. zu reagieren. Es braucht sehr viel Überwindung die eigenen antrainierten sprachlichen Normen und Reflexe auszublenden und eine geeignete Testbeziehung zu erzeugen, ohne die ProbandInnen vor den Kopf zu stoßen oder aus sozialgesellschaftlichen Gesichtspunkten unhöflich zu sein.

Die Empfehlung von Boren & Ramey, solche Thinking Aloud Test Situationen vorab zu üben wurde im Zuge der Tests für diese Diplomarbeit bestätigt. Das gesamte Forschungsteam führte zwar jeden Test von der ersten bis zur letzten Testperson professionell und motiviert durch, es konnte aber mit jedem neuen Test erkannt werden, dass die Testvorbereitung und die Testdurchführung im Forschungsteam mit jedem Mal intuitiver und routinierter abliefen. Abschließend kann bestätigt werden wie in Boren & Ramey (2000) beschrieben, dass das Thinking Aloud Protokoll nach Ericsson & Simon (1993) sich nicht für Usability Tests von Software oder deren Prototypen eignet, da es keinerlei Interaktion oder Hilfestellung mit oder seitens der ModeratorInnen erlaubt. Erst die freiere Gestaltung des weiterentwickelten Thinking Aloud Protokolls nach Boren & Ramey (2000) ermöglicht auf Basis der Sprachkommunikation ein gezieltes Suchen nach Usability Fehlern. Die beiden Autoren beschrieben die Unerlässlichkeit für Programme oder deren Prototypen, dass ModeratorInnen den ProbandInnen bei unerwarteten Eventualitäten, wie Softwarefehler oder Programmabstürze, Hilfestellungen geben. Die Notwendigkeit dieses Aspekts bestätigten die Tests an TimeRider, da Hilfestellungen bei Programmfehlern (insgesamt 76 bei den TimeRider Tests) und Programmabstürzen (insgesamt 7 bei den TimeRider Tests), im Thinking Aloud Protokoll nach Ericsson & Simon (1993) nicht vorgesehen sind.

### 8.5.4 Usability/Utility Kategorien nach Forsell & Johansson (2010)

Die Gesamtanzahl der Usability Fehler belief sich auf 1.067 Fehler welche nach Forsell & Johansson (2010) kategorisiert werden konnten. Diese Kategorien wurden zwar ursprünglich für heuristische Auswertungen erstellt, jedoch war es dadurch möglich die Ereignisse und Aktivitäten zu kategorisieren, welche während des Thinking Aloud Tests gemessen wurden. Wie zum Beispiel die Kategorisierung der Aktion, wenn UserInnen die Animation mittels direkter Manipulation der Zeitachse vor- und zurücklaufen ließen oder die Achsenparameter während eines Tasks plötzlich änderten um einen anderen Blickwinkel zu bekommen. Da die exakte Zuweisung einer Kategorie pro Fehler sehr herausfordernd war, wurde innerhalb des Forschungsteams beraten und abgestimmt, um einen Konsens pro Fehler zu finden. Durch diese Systematik konnte die Forsell & Johansson (2010) Kategorisierung auf 187 Fehler berechnet werden.

Tabelle 8: Ranking der gemessenen Usability Fehler nach Forsell & Johansson, 2010

<b>Kategorien nach Forsell &amp; Johansson (2010)</b>		
<b>Alle ProbandInnen</b>		
<b>Forsell &amp; Johansson</b>	<b>Codierung</b>	<b>Anzahl</b>
<b>spatial organization</b>	<b>B3</b>	<b>84</b>
<b>consistency</b>	<b>E16</b>	<b>58</b>
<b>prompting</b>	<b>E1</b>	<b>29</b>
<b>flexibility</b>	<b>E11</b>	<b>5</b>
<b>recognition rather than recall</b>	<b>C6</b>	<b>4</b>
<b>minimal actions</b>	<b>E7</b>	<b>2</b>
<b>orientation and help</b>	<b>B7</b>	<b>2</b>
<b>information coding</b>	<b>B5</b>	<b>1</b>
<b>remove the extraneous</b>	<b>D10</b>	<b>1</b>
<b>data set reduction</b>	<b>B9</b>	<b>1</b>

Aus der Tabelle 8 ist ersichtlich, dass vor allem drei Fehlerkategorien signifikant höher sind als die restlichen. An erster Stelle steht die Kategorie der räumlichen Darstellung (B3 = *spatial organization*). Nach Forsell & Johansson (2010) berücksichtigt diese Kategorie die Orientierung der BenutzerInnen im Informationsraum, die Verteilung von Elementen im Layout, Präzision und Leserlichkeit, effiziente Raumnutzung und Verzerrungen von visuellen Elementen. Die ProbandInnen hatten anfangs Schwierigkeiten mit der räumlichen Darstellung der Scatter Plots und mit deren Farbsättigung die je nach Zeitpunkt heller oder dunkler wurde. Dies ist auf die innovative und neuartige Darstellung der Animated Scatter Plots zurückzuführen. Die ProbandInnen haben das Programm TimeRider zuvor noch nie gesehen oder mit ihm gearbeitet und somit war auch diese Art der Visualisierung von Patientendaten aus Kohortenstudien nicht bekannt. Die Abbildung 34 zeigt zwei Cluster

(Nummer 1 in Rot und Nummer 2 in Gelb) und einen Ausreißer (Nummer 3 in Blau). Da sich bei der Interaktion mit dem Schieberegler für die Zeit die Datenpunkte bewegen und erst nach und nach angezeigt werden (je nach Messzeitpunkt), daher können sich die Cluster verschieben, sich in mehrere kleine Gruppen teilen oder über die Zeit vollständig aufgelöst werden (siehe hierfür Abbildung 35).

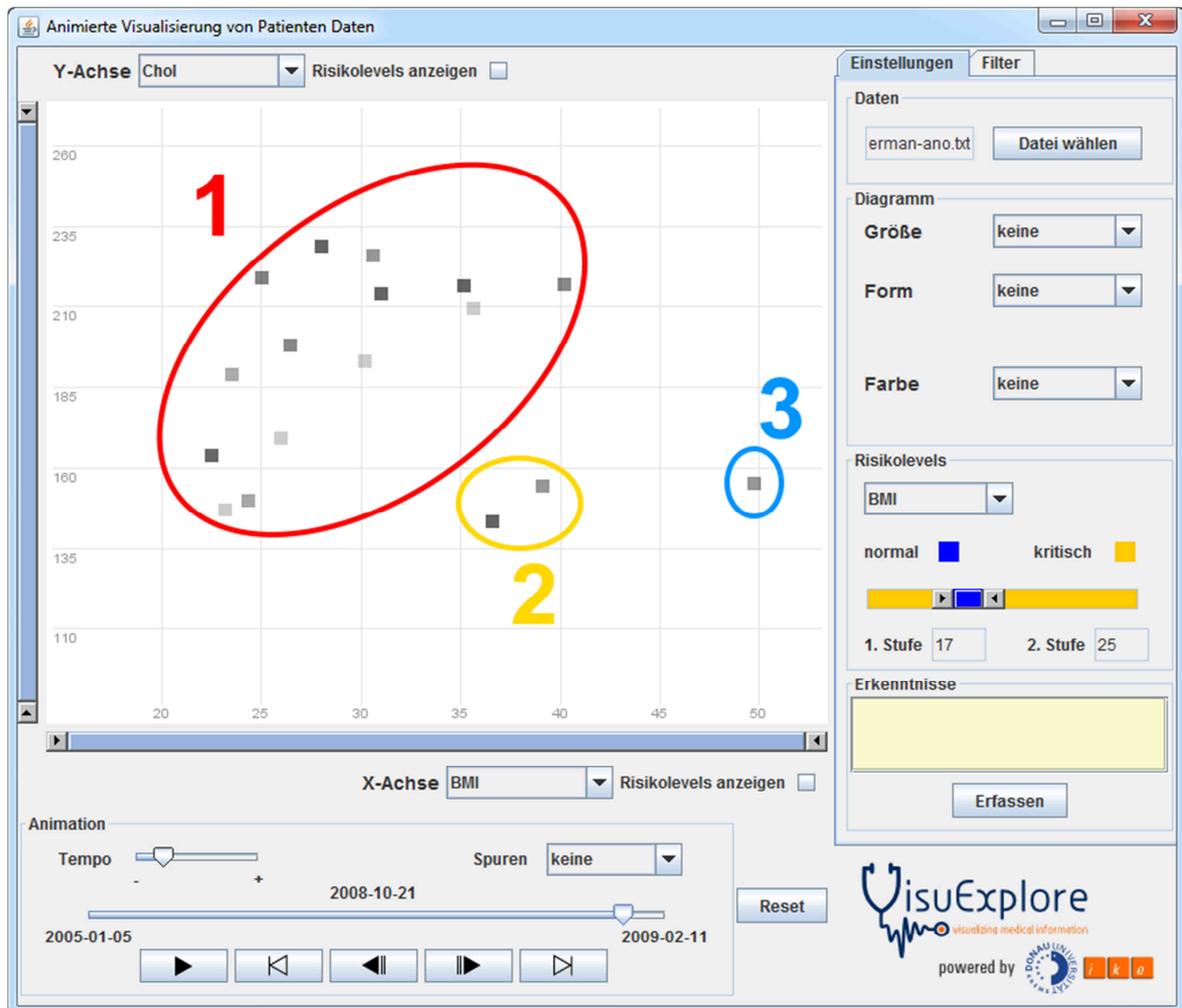


Abbildung 34: Räumliche Anordnung der Streudiagrammdaten<sup>24</sup>

Diese interaktive dynamische Anzeige, war anfänglich nicht für jeden der ProbandInnen intuitiv zu verstehen. Während der Einarbeitungsphase, spätestens jedoch nach drei Aufgabenstellungen hatten alle ProbandInnen die räumliche Organisation der Daten verstanden und konnten ohne Schwierigkeiten mit dem Programm selbständig arbeiten. Eine der Testpersonen konnte z.B. ohne aktivierte Spuren, bei der reinen Scatter Plot-Darstellung ohne Risikolevels während der Interaktion mit der Zeitachse, zuerst keinen Trend erkennen und sagte sogar, „[...] Damit kann ich wenig anfangen [...]“. Erst nachdem an die Möglichkeit von Risikolevels hingewiesen wurde, versuchte die Testperson die Aktivierungsmöglichkeit der Risikolevels aus der Einschulung sich in Erinnerung zu rufen. Nach Aktivierung

<sup>24</sup> Die bunten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen zwei verschiedene Cluster (1 und 2) in Rot und Gelb eingekreist und einen Ausreißer (3) im blauen Kreis.

der Risikolevel-Funktion und einer kurzen Interaktion mit der Zeitachse, konnte die Testperson nach einer erneuten Aufforderung des Moderators, mit dem Mauszeiger die beiden Cluster zeigen und dazu deren Trend erklären.



Abbildung 35: Clusterbewegung in TimeRider<sup>25</sup>

Die zweithäufigsten Usability Fehler wurden bei der Konsistenz des User Interface Designs (E16 = *consistency*) gemessen (siehe Abbildung 36). Laut Forsell & Johansson (2010) sollte das Design bei ähnlichen Zusammenhängen auch gleich sein (Beständigkeit des Frontenddesigns). Es stellte sich heraus, dass die Kontextmenüs der Einstellungen und Filter und damit die Einschränkung der Datenauswahl, zu längeren Entdeckungs- und Lernphasen führten. Die ProbandInnen bemängelten, dass dieses Funktionen nach keinen Ihnen bekannten Normen (z.B. MS Office, MS Windows, ...) designt waren. Des Weiteren gab es keine Menüleiste wie im MS Internetexplorer, worin einige Standardoptionen enthalten sein könnten (z.B. Datei, Ansicht, Extras, Hilfe). Es konnte beobachtet werden, dass nach den ersten drei Tasks zwar jed(e) der ProbandInnen mit diesem Fehler umgehen

<sup>25</sup> Die roten Zahlen wurden vom Autor dieser Arbeit nachträglich zur Orientierung eingefügt und zeigen die Bewegung der Cluster über die Zeit. Bild 1 (Messzeitpunkt 01.08.2007), Bild 2 (Messzeitpunkt 24.08.2007), Bild 3 (Messzeitpunkt 26.09.2007) und Bild 4 (30.10.2007).

konnte, aber auch jed(e) den Wunsch äußerte, dass eine alphabetische Sortierung in den Menüs durchgeführt werden möge. Dennoch bemängelten 4 von 10 ProbandInnen bis zum Ende der letzten Aufgabenstellung dieses unbekannte und ungewohnte Design.

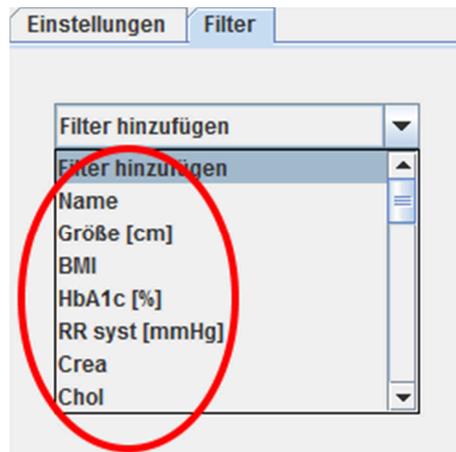


Abbildung 36: Variablen sind nicht sortiert im Kontextmenü des Filters<sup>26</sup>

Die dritthöchste Fehlerquelle waren Funktionen die nicht auffordernd genug dargestellt wurden und unklare Folgen oder Wirkungen hatten (E1 = *prompting*). In diese Kategorie fallen die beiden TimeRider Programmfunktionen, die Transparenz der Scatter Plots und die Spuren (*traces*). Beide Funktionen dienen der Darstellung des zeitlichen Verlaufs der Daten direkt im Scatter Plot. Beide Funktionen waren den ProbandInnen unbekannt und mussten erst entdeckt und verstanden werden, da diese ebenfalls nicht intuitiv verständlich waren. An dieser Fehlerquelle sieht man gut, dass die Kategorisierung schwer fällt, denn dieser Usability Fehler passt sowohl zu E16 (*consistency* = Beständigkeit des Frontenddesigns) als auch zu E1 (*prompting* = selbsterklärende Funktion). Abbildung 37 zeigt die Eigenheit dieser Funktion. In Bild 1 ist der Ausgangspunkt zu sehen, nachdem der Karteireiter Filter angeklickt und eine Filteroption aus dem zugehörigen dropdown-Menü ausgewählt wurde. Nur durch das Anklicken mittels Maustaste und gedrückt halten, lässt sich jeweils eines der beiden Rechtecke ziehen, siehe roter Kreis in Bild 3 und Bild 4 der Abbildung 37. Dadurch konnte der Wertebereich der Filterfunktion erst verändert werden. Beim Anklicken mittels Maustaste verändert sich nichts, nur wenn der Mauszeiger über dem Schiebebalken stand, dann erschien nach einigen Sekunden ein kleines Kontext-Pop-up mit dem Inhalt „Bereich einstellen“, siehe Bild 2 der Abbildung 37. Dies war besonders verwirrend, da der Schiebebalken selbst nicht einfach verändert werden konnte. Es wurde ebenfalls bemängelt, dass die Werte nicht direkt nachdem Klicken in den Wertebereich „Min“ oder „Max“ z.B. über die Tastatur manipuliert werden konnten, siehe gelber Kreis in Bild 1 der Abbildung 37. Erst nachdem die Testpersonen die Funktionen verstanden hatten und die Benutzung dieser Funktionen erlernt waren, war ein intuitives arbeiten damit ohne weitere Probleme möglich.

<sup>26</sup> Der rote Kreis wurde nachträglich vom Autor dieser Arbeit hinzugefügt und zeigt das Filtermenü mit den unsortierten Optionen, welche zur Auswahl standen.

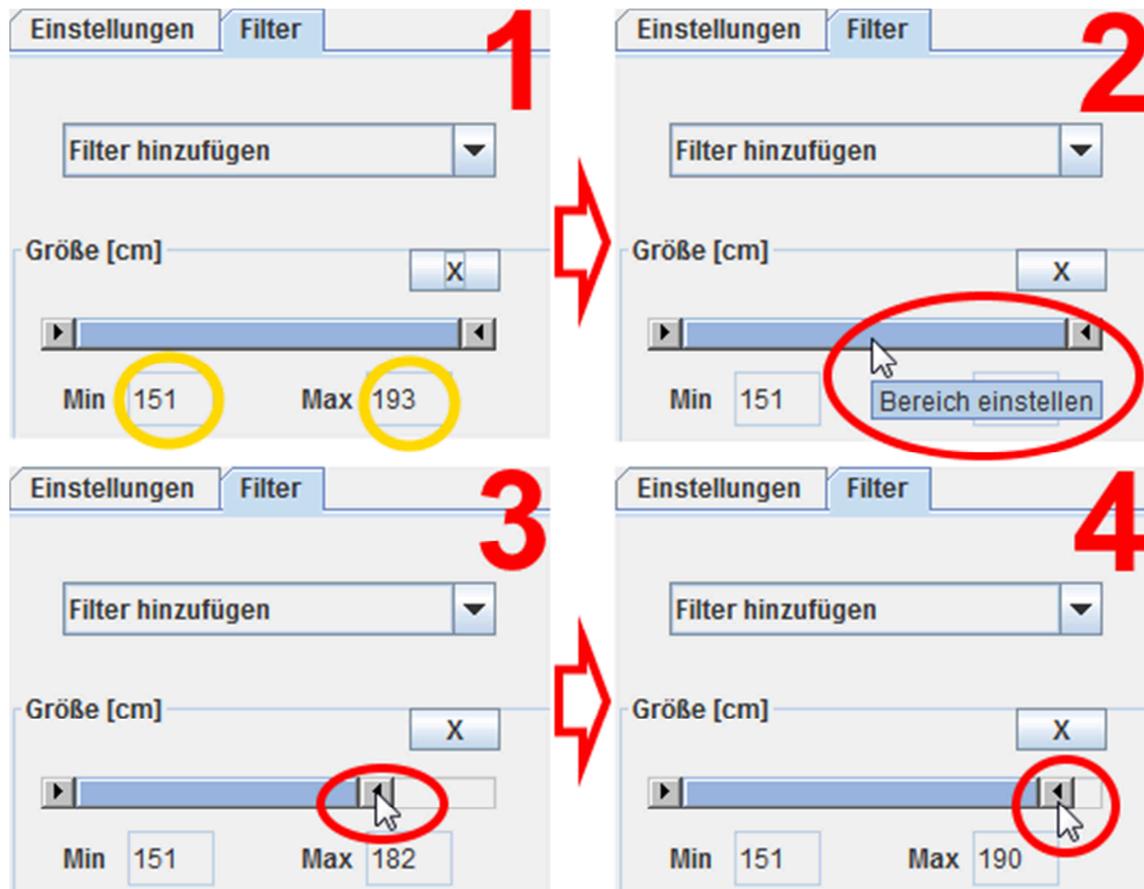


Abbildung 37: prompting-Fehler in Usability<sup>27</sup>

Somit wurden in TimeRider Schieberegler implementiert, welche weder einer allgemein verwendeten Norm (z.B. MS Office, MS Windows, ...) entsprachen, noch konnten diese intuitiv bedient werden. Die Funktion musste erst umständlich erlernt werden und bei 7 von 10 ProbandInnen musste vom Moderator aktiv eine Hilfestellung angeboten werden, damit die Funktion verstanden wurde. Nur eine der zehn Testpersonen entdeckte die Veränderung der Schieberegler des Filters >>\NBZ [mg/dl]<< des Karteireiters >>Filter<< in TimeRider zufällig. Eine andere Testperson meinte „[...] das ist ja selten dämlich [...]“ und eine weitere meinte „[...] so etwas habe ich überhaupt noch nie gesehen [...]“. Erst nach einigen motivierenden Worten und Aufforderungen seitens des Moderators, einfach die Funktion zu testen und auszuprobieren, konnten die ProbandInnen davon überzeugt werden sich trotz der anfänglichen Ablehnung mit der Funktion intensiver auseinanderzusetzen. Eine weitere Testperson beschreibt die Schiebereglerfunktion „als unlogisch“ und wünscht sich gleichzeitig die Umsetzung eines neuen Designs mit einem Schieberegler nach Vorbild der Zeitachse (siehe Abbildung 38).

<sup>27</sup> Die roten und gelben Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die komplexe Verwendung der einfachen Funktion von Schieberegler in TimeRider. Die Aktivierung der Funktion wird durch den Ablauf der Bilder 1 bis 4 dargestellt.

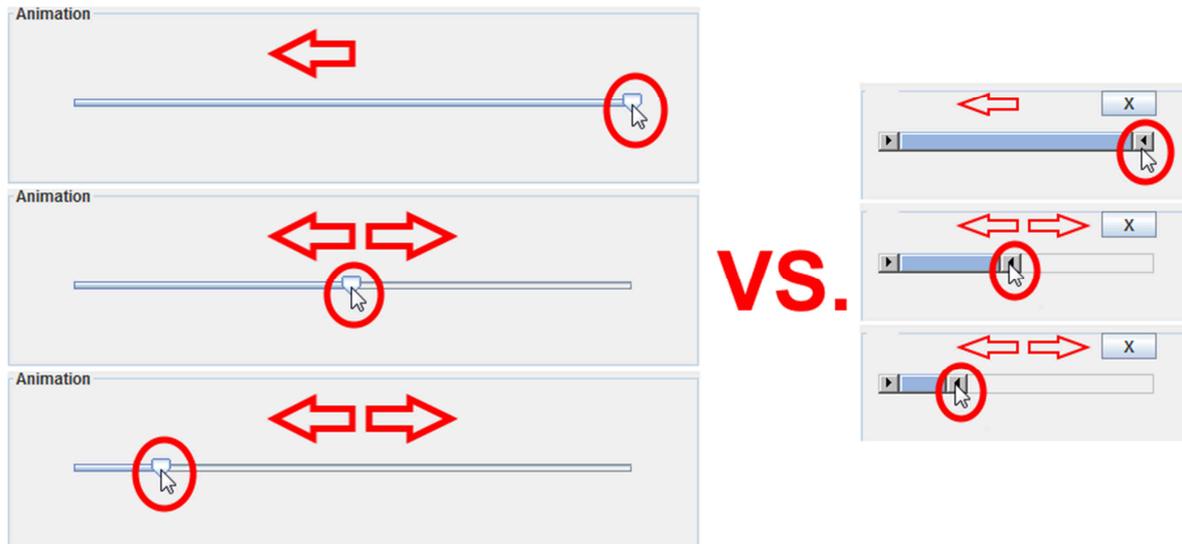


Abbildung 38: Schieberegler-Funktion der Zeitachse in TimeRider<sup>28</sup>

In Abbildung 39 sieht man die Programmfunktion der Transparenz, anhand eines Beispiels aus TimeRider. Diese Funktion soll den Abstand zum letzten Messdatum (Datum an dem eine Messung am Patienten vorgenommen wurde, z.B.: Blutabnahme und Auswertung) auf einen Blick erkenntlich machen. Während die UserInnen mit dem Schieberegler der Zeitleiste oder mit der Media Player Funktion „Play“ interagiert hat, sieht man den Transparenzeffekt der dargestellten Datenpunkte. Ist ein Messdatum auf der Zeitleiste erreicht, wird zu diesem Datum der Datenpunkt vollständig automatisch schwarz eingefärbt angezeigt, siehe dazu Bild 2 oder Bild 4 der Abbildung 39. Läuft die Zeit weiter in die Zukunft und somit vom vorherigen Messdatum weiter weg in die Zukunft, dann bewegt sich danach der Datenpunkt vom Messdatum weg und zum nächsten hin. Zwischen einem Messzeitpunkt und dem darauffolgenden Messzeitpunkt bewegt sich der Datenpunkt von einer Position zur nächsten auf einer interpolierten und unsichtbaren Bahn bis das nächste Messdatum erreicht wurde. In Bild 3 der Abbildung 39 ist der Datenpunkt nur mehr hellgrau. Diese Farbänderung oder Transparenz soll darauf hinweisen, dass sich der Datenpunkt vom Messdatum wegbewegt hat. Somit gilt: je heller ein Datenpunkt desto weiter weg ist er vom letzten Messdatum und desto näher ist er am nächsten dran, siehe hierfür Bild 1 und Bild 3 der Abbildung 39. In Bild 4 der Abbildung 39 hingegen sieht man wieder einen völlig schwarzen Datenpunkt, da wieder ein neues Messdatum erreicht wurde. Diese Visualisierung der Datenpunkte, in der sich die Transparenz über die Zeit verändert (*fading-Effekt*), wurde nur von 2 von 10 ProbandInnen direkt angesprochen. Die eine Testperson wurde auf deren ausdrücklichen Wunsch nach Hilfestellung, vom Moderator unterstützt. Die andere Testperson klärte selbst Ihre Frage, welcher Zweck die Farbänderung hat, indem sie die Datenpunkte zu unterschiedlichen Zeitpunkten mittels Kontextmenü untersuchte und dadurch den Zeitlichen Ablauf der Messdaten erkannte.

<sup>28</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die Gegenüberstellung der Schieberegler der Zeitachse (linkes Bild) und die Schieberegler der Filterfunktion (rechtes Bild).

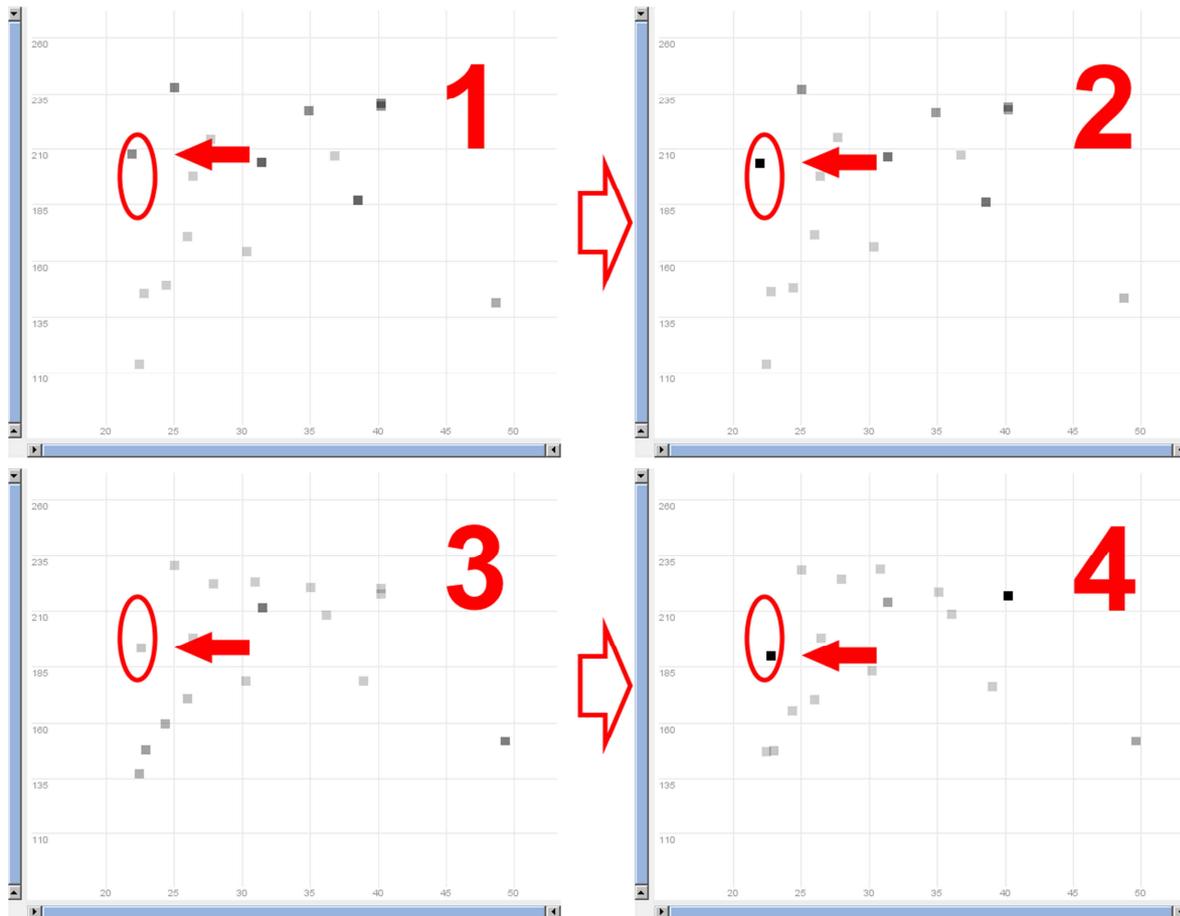


Abbildung 39: Fading Effekt der Scatter Plots in TimeRider<sup>29</sup>

### 8.5.5 Interaktionskategorien nach Yi, et al. (2007)

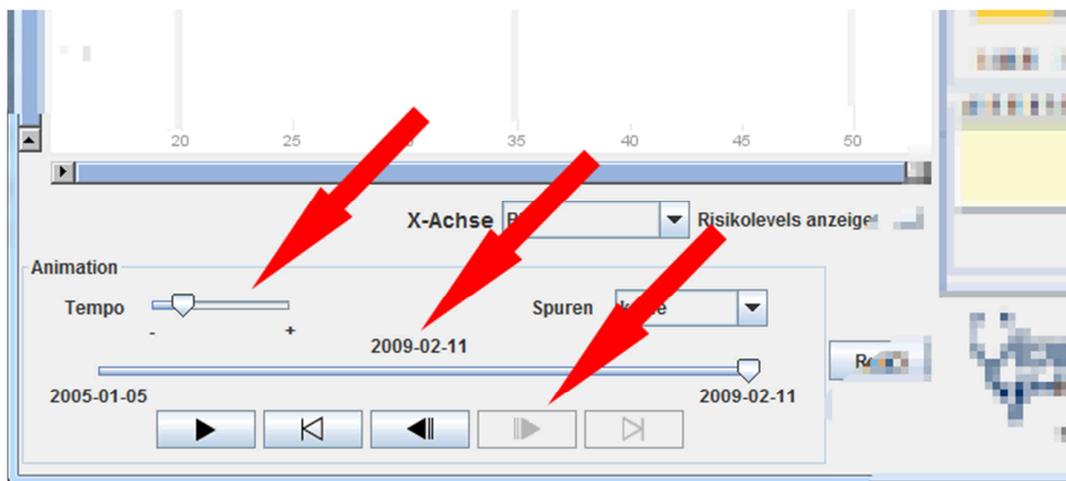
Insgesamt wurden während des Thinking Aloud Tests von TimeRider, 3.700 Interaktionen dokumentiert und nach Yi, et al. (2007) kategorisiert (siehe Kapitel 5.6). Die Kategorisierung von Aktivitäten anhand von Kategorien war jedoch herausfordernd und stellenweise schwierig, da es sich um zeitorientierte Daten handelte, welches berücksichtigt werden musste. Das Element Zeit könnte in einer Usability Analyse durch zwei Strategien umgesetzt werden: Entweder könnten die originalen Kategorien nach Yi, et al. (2007) verwendet werden und mit entsprechenden Kennzeichnungen versehen werden, dass diese zeitorientiert waren (z.B. Zeitachsen-Bewegung), oder es könnte eine weitere Kategorie für Zeit hinzugefügt werden, wie es in dieser Masterarbeit mit der Kategorie „Zeitachse“ durchgeführt wurde (siehe Tabelle 9). Diese Kategorie „Zeitachse“ umfasst alle Aktionen der UserInnen, welche mit der Manipulation des Ablaufs der Zeit zu tun hatten, wie Einstellen des Wiedergabetempos und jegliche Interaktion mit der Zeitachse und deren Media Player Funktion (siehe Abbildung 40).

<sup>29</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die Änderung der Farbtransparenz der Datenpunkte im Verhältnis zum letzten Messdatum.

Tabelle 9: Auswertung der Interaktionen nach Yi, et al. (2007)

Kategorien nach Yi, et al. (2007)	
Alle ProbandInnen	
Yi, et al.	Anzahl
abstract/ elaborate	1201
Variablen	919
Zeitachse	598
filter	277
encode	276
reconfigure	244
connect	135
explore	49
select	-

Des Weiteren wurde eine weitere Kategorie hinzugefügt namens „Variablen“, in diese fallen alle Aktionen der UserInnen, bei denen sich die ProbandInnen die Variablen genauer ansahen. Diese Kategorie gilt sowohl für das „dropdown“ Menü der kartesischen Achsen, als auch für das „dropdown“ Menü der Einstellungen „Diagramm“ für Größe, Form und Farbe, sowie den Inhalt des „mouseover-tooltip“.

Abbildung 40: Betroffene Funktionen der zusätzlichen Kategorie „Zeitachse“<sup>30</sup>

Es folgt nun das Ranking der häufigsten Interaktionen der UserInnen, kategorisiert nach Yi, et al. (2007):

1. Die häufigste Interaktion der UserInnen war mit eindeutigen Abstand „Abstract/Elaborate“ (zeig mir mehr oder weniger Details). Laut Yi, et al. (2007) beinhaltet diese Interaktionskategorie jede Aktion der BenutzerInnen, die eine Möglichkeit bietet den Grad der Abstraktion der Datendarstellung anzupassen. Ein Beispiel für eine „Abstracte/Elaborate“-Technik sind die einfachen Tool-Tipp In-

<sup>30</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die betroffenen Funktionen in TimeRider (Temposchiebereger, Zeitschiebereger, Media Player Funktionen), welche in die zusätzliche Kategorie „Zeitachse“ fallen.

teraktionstechniken welche Detailinformationen anzeigen, wenn der Mauszeiger über Datenpunkte bewegt wird, siehe Abbildung 41. Dies war auch in TimeRider die häufigste Funktion, welche mittels „*mouseover/hovering*“-Kontextmenü der Datenpunkte in TimeRider implementiert wurde. Da dieses Kontextmenü alle gesammelten Daten der jeweiligen PatientInnen enthielt, waren die ÄrztInnen von TimeRider eindeutig verleitet, sehr lange und sehr oft dieses Kontextmenü zu öffnen und Informationen zu betrachten. Obwohl das klassische Beispiel für diese Interaktionskategorie das „Zoomen“ ist, wurde es durchschnittlich in nur einem Viertel (ca. 300 von 1.201 Interaktionen) aller Aktionen der UserInnen in dieser Kategorie gesamt eingesetzt.

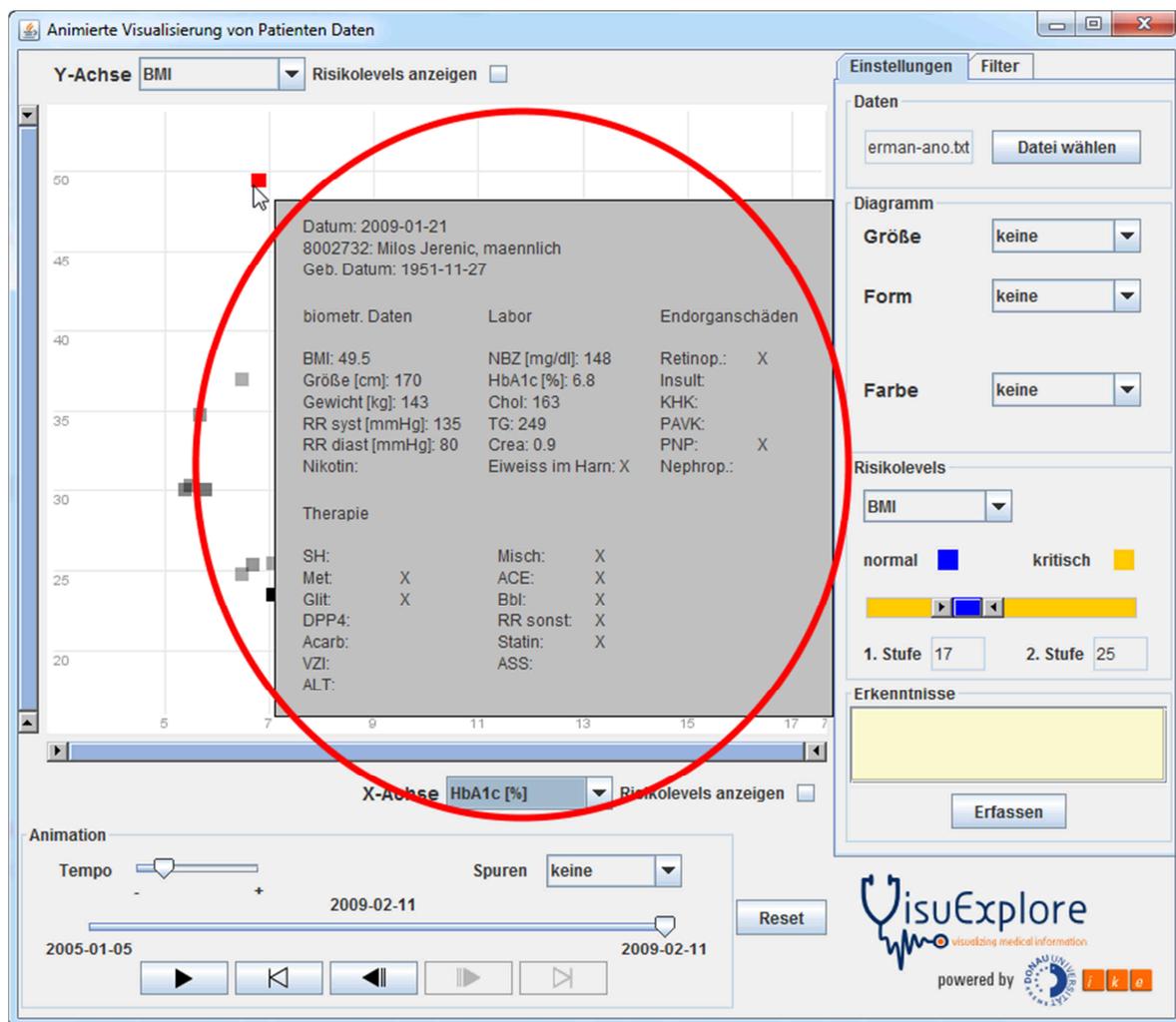


Abbildung 41: Häufigste Funktion der Kategorie „Abstract/Elaborate“<sup>31</sup>

2. An zweiter Stelle der UserInnen-Interaktionen schaffte es die zusätzliche Kategorie „Variablen“. Die Häufigkeit dieser Interaktion, ergab sich aus der Ausbildung der ÄrztInnen. Sie werden schon während des Studiums auf diverse Statistiken und

<sup>31</sup> Die rote Kennzeichnung wurde vom Autor dieser Arbeit nachträglich eingefügt und zeigt die häufigste Funktion „*mouseover/hovering*“ in der Kategorie „Abstract/Elaborate“.

- Grenzwerte, -bereiche trainiert. Diese Daten können auf einen Blick im Kontextmenü eines Datenpunkts zum letzten jeweiligen Messdatum abgelesen werden.
3. An dritter Stelle befand sich die Interaktion mit der Zeitachse, welche durch die neu hinzugefügte Kategorie „Zeitachse“ gemessen wurde. Nachdem die ProbandInnen die Funktion der Zeitachse entdeckt hatten, wurde diese auch gerne genutzt. Am meisten wurde der Schieberegler der Zeitachse verwendet um die Animation der Datenpunkte des Scatter Plots mittels eigener Geschwindigkeit vor und zurück über die Zeit ablaufen zu lassen. Die Media Player Funktion wurde von nur 4 von 10 ProbandInnen ausprobiert, wobei eine Aussage einer Testperson die durchgängige Meinung treffend zusammenfasste, dass „[...]die Interaktion mit dem Schieberegler der Zeitachse ist gegenüber der Videoplayer-Funktion "Play" viel effektiver [...]“ (siehe Abbildung 28).
  4. Die Filterkategorie (zeig mir etwas unter gewissen Bedingungen) befindet sich auf Platz vier der kategorisierten Interaktionen der UserInnen. Diese Funktion bereitete einige Schwierigkeiten durch die unsortierten Auswahlmenüs. Daher konnte beobachtet werden, dass die Filterfunktionen von nur 3 von 10 Testpersonen öfters als in den Aufgabenstellungen vorgesehen verwendet wurden. Die meisten ProbandInnen mieden die Filterfunktionen, da sie laut einer Testperson „[...] noch nicht ganz ausgereift scheinen [...]“.
  5. An fünfter Stelle befindet sich die „Encode“ Kategorie (zeig mir eine andere Darstellung) nach Yi, et al. (2007). Diese wurde mittels einer Anpassungsoption für die Größe, Farbe und oder Form der Datenpunkte im Scatter Plot in TimeRider implementiert. Alle ProbandInnen verwendeten diese Funktion, aber erst nach der zweiten Aufgabenstellung, da dieses es zwingend verlangte. Danach verwendeten immerhin 6 von 10 ProbandInnen diese Anpassungsoption selbständig, obwohl dies in den Aufgabenstellungen nicht mehr zwingend erforderlich war.
  6. Die „Reconfigure“ Kategorie (zeig mir eine anderes Arrangement) nach Yi, et al. (2007) ist auf Platz sechs zu finden. In diese Kategorie passte vor allem die Veränderung der Achsenparameter. Da die ersten drei Aufgabenstellungen bestimmte Parameter vorgaben, konnten die ProbandInnen nicht viel experimentieren. In der vierten Aufgabenstellung, gab es jedoch kein Halten mehr und alle ProbandInnen veränderten nach eigenen Vorstellungen die Achsenparameter um neue Erkenntnisse zu gewinnen oder einfach verschiedenste Arrangements zu vergleichen und auszuprobieren.
  7. Auf der siebten Stelle befindet sich die „Connect“ Kategorie (zeig mir verwandte Datenpunkte) nach Yi, et al. (2007). Diese Interaktionskategorie wurde mittels

Traces (Spuren) umgesetzt. Diese Funktion wurde teilweise verhalten aufgenommen. Auf den ersten Blick oder bei einem eingeschränkten Datensatz wurden die Spuren als sehr hilfreich beschrieben, da es noch übersichtlich war und auf einen Blick erkannt werden konnte, wie der bisherige Verlauf der Werte war. Als jedoch mehrere Datenpunkte mit Spuren angezeigt wurden kreuzten oder überlappten die Spuren sich gegenseitig. Dadurch empfanden dann 7 von 10 ProbandInnen die Spuren als zu unübersichtlich und alle ProbandInnen waren sich einig „[...] die Traces sind [...] nicht für große Datensätze geeignet, da sehe ich keine Sinnhaftigkeit [...]“.

8. Die vorletzte Kategorie ist die „*Explore*“ Kategorie (zeige mir etwas anderes) nach Yi, et al. (2007). Das treffendste Beispiel, wie in Kapitel 5.6 beschrieben ist das „panning“. Diese Funktion war in TimeRider implementiert, da sowohl das Panning im Scatter Plot im hineingezoomten Zustand möglich war, als auch bei jedem Schieberegler und sogar bei der problematischen Schiebebalkenfunktion (siehe Kapitel 8.5.4).
9. Die *Select*-Kategorie (markiere etwas als interessant) befindet sich am letzten Platz, da sie zwar zu den sieben Kategorien von Yi, et al. (2007) gehört, aber in keiner Art und Weise in TimeRider implementiert wurde.

Es folgen nun Hinweise der ProbandInnen zu Fehlern oder Designschwierigkeiten von TimeRider, welche während des Thinking Aloud Tests dokumentiert wurden.

### 8.5.6 Dokumentierte Hinweise der ProbandInnen

Alle TeilnehmerInnen konnten zwar Trends in dem Verhalten der Punktwolken über die Zeit, deren Muster und Zusammenhänge in den Daten erkennen, jede einzelne Testperson hatte jedoch, trotz der kurzen Einschulungsphase am Anfang, Schwierigkeiten mit der Bedienung oder der Navigation des User Interface von TimeRider. Es ist jedoch zu bedenken, dass dieses Programm vor den Tests den ProbandInnen völlig unbekannt war und sie es beim Test das erste Mal verwendeten. Nachfolgenden sind nun die wichtigsten Kritikpunkte der ProbandInnen aufgelistet, welche sich aus den aufgezeichneten Thinking Aloud Protokollen herauskristallisiert haben:

1. **Achsenbezeichnung:** Die Bezeichnung der Achsen sollte verbessert werden, die ProbandInnen verwechselten die Achsen aufgrund der unklaren Beschriftung (siehe Abbildung 42).

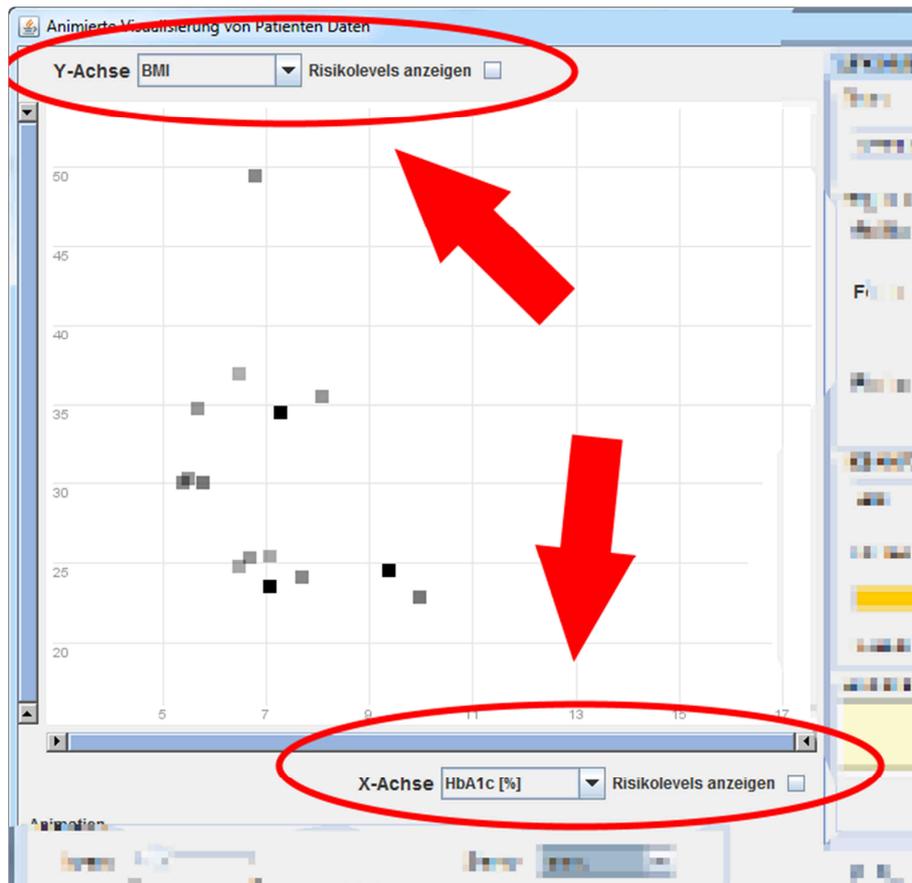


Abbildung 42: Verwirrende Achsenbezeichnung in TimeRider<sup>32</sup>

2. **Kontextmenü 1:** Alle ProbandInnen beklagten, dass die „dropdown“ Menüs der Filter zumindest alphabetische geordnet sein sollten. Es wurde auch der Wunsch geäußert, weitere Individualisierungen der Ordnungsmöglichkeiten anzubieten (z.B. Sortierung von A nach Z und Z nach A, usw.) (siehe Abbildung 43).

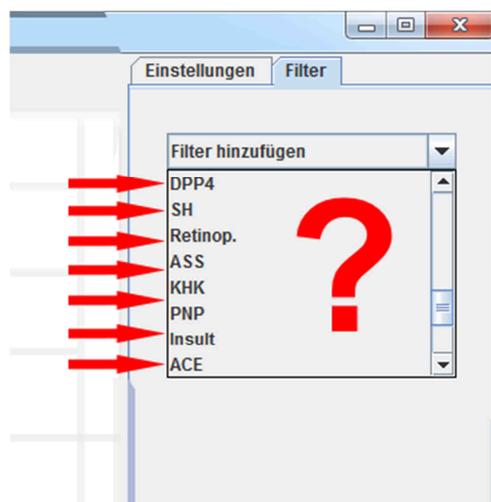


Abbildung 43: Unsortiertes Filtermenü in TimeRider<sup>33</sup>

<sup>32</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die Achsenbeschriftungen, welche von den ProbandInnen als verwirrend empfunden wurden.

<sup>33</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen das unsortierte dropdown Menü des Karteireiters „Filter“.

3. **Kontextmenü 2:** Alle ProbandInnen bemängelten, dass die „dropdown“ Menüs der Filter nicht mittels Keyboard manipulierbar waren.
4. **Kontextmenü 3:** Des Weiteren wurde bemängelt, dass die Funktion zur Öffnung des Kontextmenüs, mittels Anklicken eines einzelnen Patientendatenpunkts durch die rechte Maustaste, nicht implementiert wurde.
5. **„Overlapping“ 1:** Bei einer Farbänderung der Datenpunkte und der Standardvorgabe der Farbänderungsdarstellung blau und gelb, wird durch den Einsatz der Transparenzvisualisierung, die gelben Punkte so hell, dass sie teilweise nicht mehr wahrnehmbar sind (siehe Abbildung 44). Des Weiteren wurde gebeten eine Individualisierungsoption zu implementieren, welche eine andere individuelle Farbcodierung ermöglicht.
6. **„Overlapping“ 2:** Die ProbandInnen wiesen darauf hin, dass bei Überlappungen, sowohl der Traces als auch der Scatter Plots, weder ersichtlich war, welcher Patientendatenpunkt oben ist und welcher unten (siehe Abbildung 45). Ebenfalls wurde darauf hingewiesen, dass bei einer Farbänderung der Datenpunkte (der Standard der Farbänderungsdarstellung ist blau und gelb) bei Überlappung die beiden Datenpunkte zu einen grün angezeigten verschmelzen (siehe Abbildung 44).

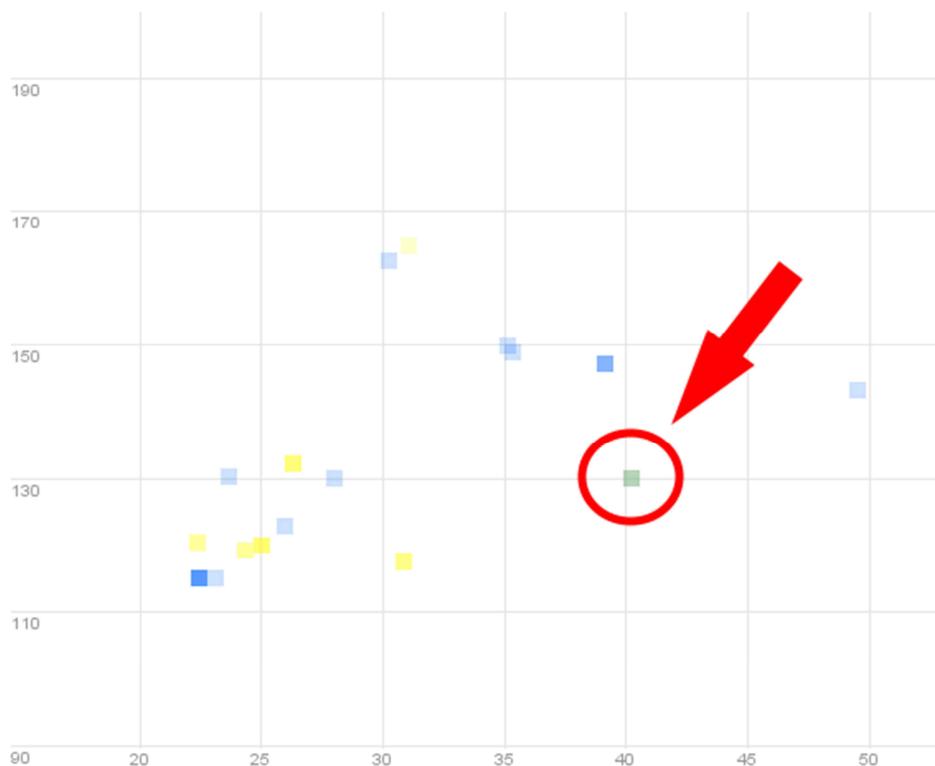


Abbildung 44: Überlagerung der Scatter Plots kann zur Verfärbung führen<sup>34</sup>

<sup>34</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen, dass es bei einer geänderten Farbkodierung der Datenpunkte bei Überlagerung zu Farbverfälschungen kommt.

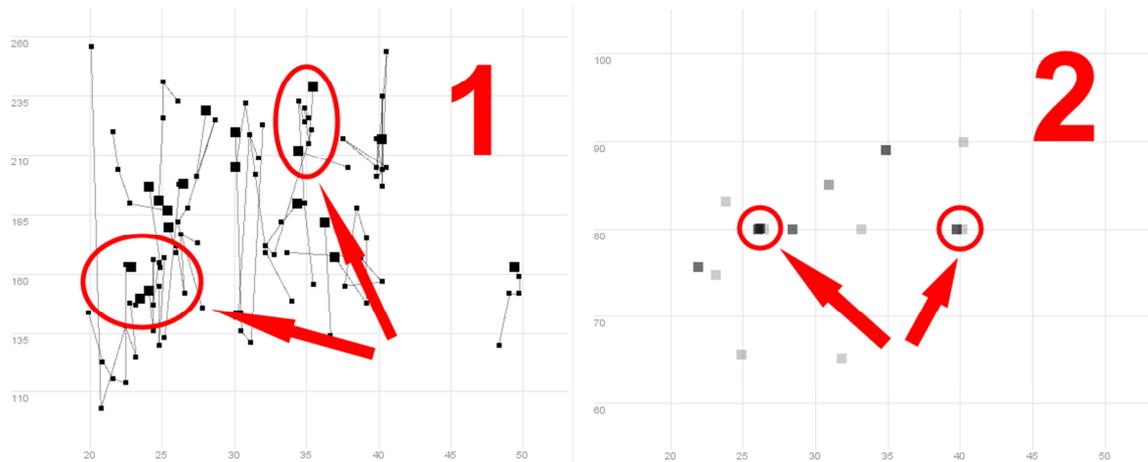


Abbildung 45: overlapping führt zu Unübersichtlichkeit<sup>35</sup>

7. **„Prompting“:** Alle ProbandInnen bemängelten, dass keine einzelnen Patientendatenpunkte markiert werden konnten und auch markiert blieben, durch etwa einfaches Anklicken mittels der linken Maustaste.
8. **Speichern:** Verschiedenste vorgenommene Einstellungen sollten gespeichert werden können, damit die benutzerspezifische Darstellung zu einem späteren Zeitpunkt, einfach abrufbar ist.
9. **Traces:** Die ProbandInnen bemängelten, dass die Spurendarstellung (*traces*) verwirrend und unverständlich ist und dass bei „nur“ 20 Patientendaten, denn meistens enthalten Kohortenstudien mehr PatientInnen und deren Daten. Ebenfalls konnte nicht eindeutig, selbst bei langsamem Zeitablauf, immer nachverfolgt werden, wie die Spuren sich entfalten.
10. **Zeitanimation:** Die Geschwindigkeit der Zeitachsen-Wiedergabefunktion ist viel zu langsam. Der Zeitschieber wurde immer manuell betätigt, dieses Verhalten der UserInnen macht jedoch die Media Player Funktionen überflüssig.
11. **Zoom:** Die ProbandInnen störte, dass während des Zoomens die Anzeige ausgeblendet wird. Es wurde bemerkt, dass dadurch der Überblick verloren geht. Eine Möglichkeit wäre es „Zoom Buttons“ einzufügen mit einer Zoom-Stufenanzeige und stufenlosen Zoom (z.B. 1% - 100% usw.).
12. **Risikolevels:** Die ProbandInnen wurden durch die Nomenklatur und die Farbgebung der Risikolevels irritiert und verwechselten die Risikobereiche mit den Normbereichen (siehe Abbildung 46). Diese Fehlinterpretation kommt zustande, da bei der Aktivierung der Risikobereiche die Normbereiche blau eingefärbt werden und somit hervorgehoben werden. Somit kommt es zu einer Verwechslung weil erwar-

<sup>35</sup> Die roten Kennzeichnungen wurden vom Autor dieser Arbeit nachträglich eingefügt und zeigen die Unübersichtlichkeit bei Überlappungen.

tet wird, dass bei Aktivierung der Risikolevels, diese auch eindeutig besser sichtbar werden und nicht die Normbereiche. Eine derartige Verwechslung ist in der Humanmedizin inakzeptabel, vor allem wenn es sich um Risikobereiche handelt. Dieser Usability Fehler muss unbedingt behoben werden!

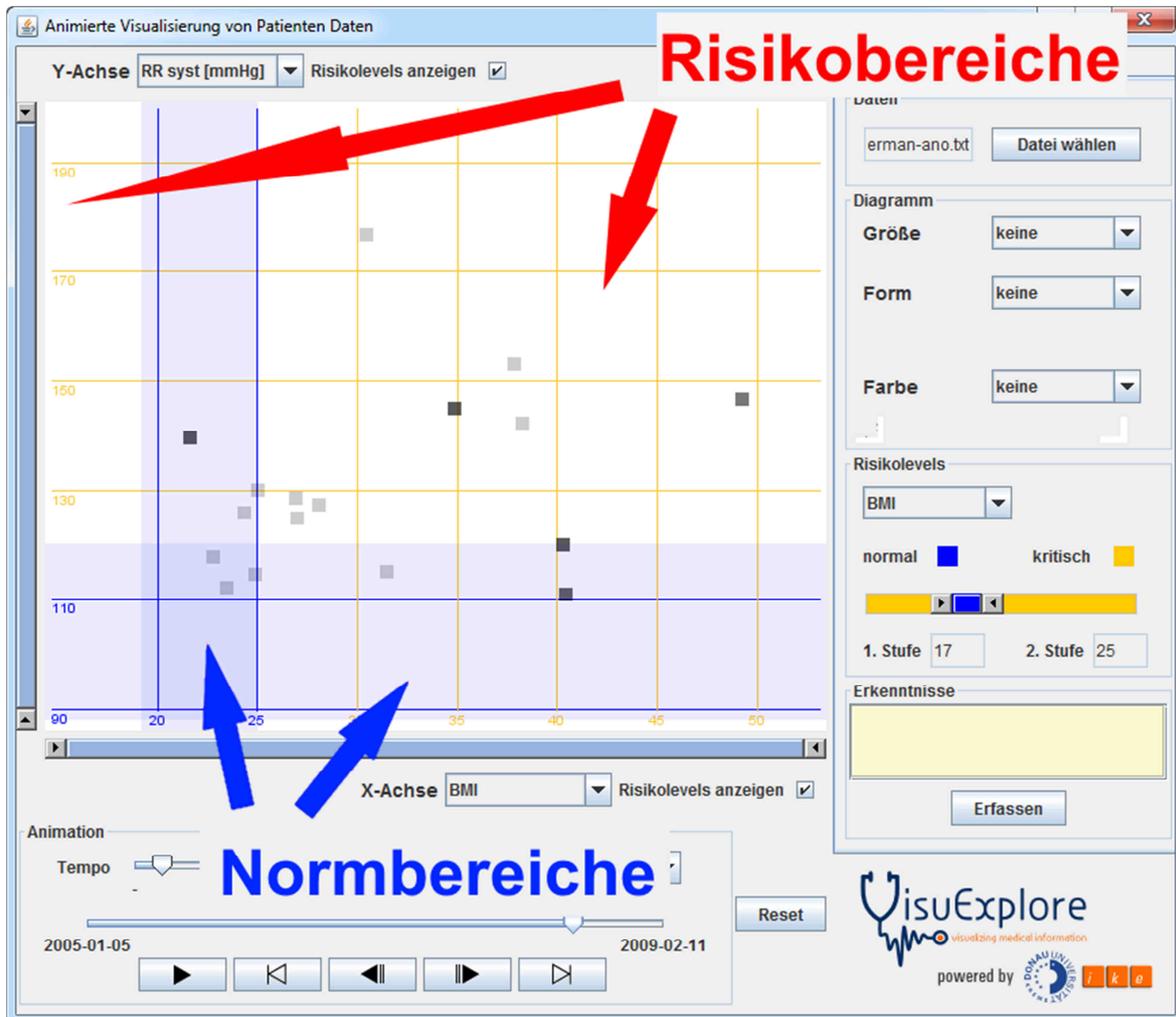


Abbildung 46: Risikolevel Aktivierung, hebt den Normbereich hervor

### 8.5.7 Abschließende Beurteilung

Die Ergebnisse dieser Untersuchung zeigen, dass die Implementierung von animierten Patientendaten einer Kohortenstudie die ÄrztInnen in deren Arbeit unterstützt und liefern somit weitere bereichsspezifische Beweise für die Diskussion über die Effektivität der Verwendung von Animationen.

Der Usability Test welche durch die Thinking Aloud Methode nach Vorbild des Boren & Ramey (2000) Protokolls durchgeführt wurde, offenbarte sowohl die Gedankenabfolge der ProbandInnen bei der Lösung der gestellten Aufgaben als auch die Usability Fehler des User Interface des Programms. Ebenfalls konnte dadurch der Lerneffekt dokumentiert

werden, der während des Tests beobachtet und mittels TASK IV durch freie Parameterwahl und ohne weitere Vorgaben überprüft wurde.

Generell war zu beobachten, dass es allen StudienteilnehmerInnen möglich war, die gestellten vier Aufgaben zu lösen. Abschließend ist noch zu sagen, dass einige Potentiale zur Verbesserung von TimeRider aufgezeigt wurden. Vor allem die Darstellungsfunktion der Risikolevels muss unbedingt überarbeitet werden, damit es bei den zukünftigen ÄrztInnen, welche mit dem Programm arbeiten werden, zu keinen Verwechslungen kommen kann. Alle weiteren Fehler und Schwierigkeiten sollten ebenfalls in ein „*Redesign*“ einfließen, wie von Nielsen (siehe Kapitel 5.3) empfohlen, damit die nächste Version von TimeRider noch intuitiver erlernbarer ist.



## 9 Zusammenfassung

Im Zuge dieser Diplomarbeit wurde ein Lerneffekt von ÄrztInnen bei der Interaktion mit dem Programm TimeRider durch vier Aufgabenstellungen überprüft. Die ersten drei Aufgaben wurden als Einarbeitungsphase eingesetzt und die letzte Aufgabenstellung überprüfte den Lerneffekt, durch eine freie Interaktion mit dem Programm, wodurch die ProbandInnen animiert waren selbständig mit dem Programm zu arbeiten. Während der Lösung der Aufgabenstellungen wurde ein Thinking Aloud Tests nach Boren & Ramey (2000) durchgeführt, bei dem rund acht Stunden Videomaterial entstanden. Die Auswertung dieser Daten ergaben 187 Fehler kategorisiert nach Forsell & Johansson (2010). Des Weiteren wurden 76 Programmfehler und fehlerhafte Funktionen entdeckt und es kam insgesamt zu sieben Programmabstürzen. Diese Daten sollen in zukünftigen Updates des Programms TimeRider berücksichtigt werden. Die ProbandInnen tätigten insgesamt 3.700 Interaktionen kategorisiert nach Yi, et al. (2007). Daraus ließ sich ein Trend erkennen, dass die ÄrztInnen sich durch die Kontextmenüs der Datenpunkte verleiten ließen, die Informationen genauer zu prüfen als einfach die Animation des Scatter Plots zu nutzen, um Erkenntnisse zu gewinnen. Alle ÄrztInnen gaben positives Feedback zur Funktionalität des Programms, auch wenn einige dieser Funktionen neu für die Testpersonen waren. Eine Testperson meinte sogar: „[...]Ansonsten finde ich das Tolle an dem Programm, dass das eigentlich [...] echt so Control Chart mäßig eben arbeitet und [...] noch nicht Signifikanzen irgendwie vorweg nimmt, sodass man da gleich einmal wirklich Trends anschauen kann [...]“ Es wurde sogar von allen ÄrztInnen empfohlen das Programm TimeRider nicht nur im klinischen Arbeitsalltag, sondern vor allem in der medizinischen Aus- und Weiterbildung einzusetzen.

### 9.1 Grenzen der Untersuchung

Die Grenzen dieser Diplomarbeit liegen nicht in der Anzahl der Testpersonen, welche am Usability Test teilnahmen (siehe Kapitel 8.3), sondern in der Länge des Tests. Es konnten die 10 Testpersonen, welche den Test zum Ersten Mal sahen, nur für kurze Zeit im Rahmen dieser Magisterarbeit damit arbeiten. Dies erklärt auch die anfänglichen Schwierigkei-

ten bei der Benutzung des Programms. Es ist daher keine allgemeingültige Aussage aus den gemessenen Daten ableitbar, welche TimeRider einen sofortigen Einsatz im klinischen Alltag attestieren könnte.

Darüber hinaus wurden nur ÄrztInnen und kein medizinisches Personal als Testpersonen herangezogen. Daher konnte das Programm auch nicht aus Sicht des medizinischen Personals einer Klinik getestet werden. Die Bedürfnisse und Lösungsstrategien dieser beiden Benutzergruppen unterscheiden sich jedoch im klinischen Alltag und damit auch die Verwendung und der Einsatz von medizinischen Programmen.

Laut Nielsen (siehe Kapitel 8.3) sind maximal 10 Testpersonen (ProbandInnen) ausreichend, denn ein mehr an Test-UserInnen ist nicht signifikant aussagekräftiger. Die Humanmedizin ist jedoch ein sehr spezielles Fachgebiet mit eigenen gesetzlichen Regeln und Vorschriften, da von den Entscheidungen der Humanmediziner nichts Geringeres als Menschenleben abhängig sind. Bevor ein Einsatz eines medizinischen Programms im klinischen Alltag möglich ist, wird es eine Herausforderung sein die Verantwortlichen davon zu überzeugen, dass ein paar Usability Tests mit maximal 10 ÄrztInnen ausreichend sind. In der Regel sind Professionals humanmedizinischen Studien gewöhnt, in denen die Devise ist, je mehr Testobjekte, desto signifikanter ist ein Test. Dies wurde auch von 7 der 10 ProbandInnen dieser Magisterarbeit bemängelt, warum nur 20 Patienten insgesamt sich im Datensatz befinden. Eine der ProbandInnen merkte sogar an, dass man zwar Cluster und Trends erkennen kann, diese aber nicht aussagekräftig sind, da das Patientenkollektiv so klein ist.

## 9.2 Ausblick

TimeRider ist ein Programm mit einer verbesserten Visualisierung von animierten Streudiagrammen (*Animated Scatter Plots*), welches die drei Herausforderungen bei der Erkundung von visuellen Patientendaten einer Kohortenstudie meistert: unregelmäßige Stichprobenziehungen, Datenverschleiß und Datensätze mit unterschiedlichen Zeitabschnitten. Diese Herausforderungen werden durch die visuelle Darstellung der Zeit mittels Transparenz und Spuren (*Traces*) bewältigt. Die Auswertung der Daten zeigte, dass die ÄrztInnen durch die Verwendung von TimeRider neue Erkenntnisse gewinnen konnten. Des Weiteren können die entdeckten Fehler und Systemfehler für ein „*Redesign*“ und eine Aktualisierung des Programms verwendet werden. Die Forschungsergebnisse lassen den Schluss zu, dass weitere Untersuchungen in Zukunft durchgeführt werden sollten, wie z.B. bei überlagerten Datenvisualisierungen bzw. Datenpunkten. Davon betroffen sind auch die in TimeRider verwendeten Spuren (*Traces*), welche sehr effektiv Verläufe darstellen können, wenn sie sich nicht gegenseitig überlagern und das Konzept von *Traces* verstanden wurde. Eine

optimale Lösung der Traces Darstellung könnte in zukünftigen Untersuchungen erforscht werden. Ebenfalls sollte ein längerer Test als in dieser Magisterarbeit durchgeführt werden, z.B. 4 Mediziner welche ein Monat lang sich mit TimeRider beschäftigen und dabei eine Thinking Aloud Methode anwenden und sich selbst aufnehmen. Es ist zu erwarten, dass das Langzeitverhalten von Test-UserInnen sich von jenen in einen kurzen Test unterscheiden, da die Lerneffekte mehr Wirkung haben werden. Die Masse an Daten würde wahrscheinlich den Rahmen einer Diplomarbeit sprengen und wäre für eine potentielle Doktorarbeit passender.

Jedes Visualisierungsprogramm sollte spezifische Anpassungen der BenutzerInnen ermöglichen um eine größtmögliche Akzeptanz unter den verschiedensten UserInnen zu erreichen. Für TimeRider ist dies in der untersuchten Version teilweise umgesetzt worden, da dieses Programm jedoch vorrangig für den humanmedizinischen Bereich eingesetzt wurde, sollten die Lösungsstrategien von ÄrztInnen berücksichtigt werden. In der Humanmedizin werden ÄrztInnen auf eine spezifische Weise geschult, Daten und vor allem Grenzwerte mittels spezieller Zusammenhänge (z.B. Zusammenhang zwischen Blutdruck und Körpergewicht eines Patienten) zu analysieren und zu interpretieren. Wenn die in dieser Magisterarbeit beschriebenen Fehler und Usability Probleme behoben werden und die speziellen Bedürfnisse von ÄrztInnen berücksichtigt und unterstützt werden, dann kann TimeRider ein Standard-Visualisierungstool in der humanmedizinischen Ausbildung und im klinischen Alltag werden, sowie eine größtmögliche Akzeptanz in der Humanmedizin erlangen.



## Bilderverzeichnis

Abbildung 1: User Interface des zu untersuchenden Programms TimeRider .....	1
Abbildung 2: Zeitlicher Ablauf von Kohortenstudien (angelehnt an Grimes & Schulz, 2002).....	9
Abbildung 3: User Interface des VisuExplore-Prototypen.....	12
Abbildung 4: Napoleons Russlandfeldzug dargestellt von Charles Joseph Minard (Homepage von „InfoVis.net“, 2015).....	13
Abbildung 5: Darstellung des Powsner & Tufte „all data chart“ (Homepage der „IFS“, 2006).....	14
Abbildung 6: Graphische Zusammenfassung des Patientenstatus (Homepage von Edward Tufte, 2015). .....	15
Abbildung 7: User Interface einer elektronischen Patientenakte (Plaisant, et al., 1998, p.3). .....	16
Abbildung 8: Bild der Höhlenmalerei in der El-Castillo-Höhle (vgl. Pike, et al., 2012, Supplementary Materials, p.14).....	20
Abbildung 9: Beispiel einer wissenschaftlichen Visualisierung (vgl. Schütz, 2013, p.448) .....	22
Abbildung 10: Beispiel für medizinische Visualisierung (vgl. Homepage der Universität zu Lübeck, 2014).....	25
Abbildung 11: Small Multiples Darstellung in einer weiterentwickelten Version von TimeRider (Neubauer, 2012, p.46).....	37
Abbildung 12: Beispielbilder von einigen zufällig gewählten Streudiagrammen (in Anlehnung an Homepage des „KSA“, 2013). .....	38
Abbildung 13: Datenpunkteschwarm in TimeRider.....	39
Abbildung 14: Animation in TimeRider .....	40
Abbildung 15: Schematische Darstellung der Usability-Wechselwirkung (in Anlehnung an Shackel & Richardson, 1991, p.24).....	42
Abbildung 16: Explore Technik in TimeRider.....	58
Abbildung 17: Reconfigure Technik in TimeRider.....	59
Abbildung 18: Encode Technik in TimeRider .....	60
Abbildung 19: Abstracte/Elaborate Technik - „mouseover“ in TimeRider .....	61
Abbildung 20: Abstracte/Elaborate-Technik - „zoomen“ in TimeRider.....	61
Abbildung 21: Filter Technik in TimeRider.....	62
Abbildung 22: Connect Technik - „traces“ in TimeRider.....	63
Abbildung 23: Connect Technik - „highligthing“ in TimeRider.....	63
Abbildung 24: behavioristisches Wahrnehmungsbild.....	66

Abbildung 25: TimeRider Benutzerschnittstelle (User Interface/Frontend/GUI).....	90
Abbildung 26: Achsen „dropdown“Menü in TimeRider .....	90
Abbildung 27: Variablendarstellung in TimeRider.....	91
Abbildung 28: „timeslider“ Funktion in TimeRider .....	92
Abbildung 29: „Risikolevels“ Funktion in TimeRider .....	93
Abbildung 30: „details on demand“ Funktion, mittels „mouseover“ in TimeRider .....	94
Abbildung 31: Small Multiples Visualisierung von TimeRider .....	95
Abbildung 32: Small Multiples Visualisierung mit Traces von TimeRider .....	96
Abbildung 33: Verhältnis zwischen gefundenen Usability Problemen und Test-UserInnen .....	97
Abbildung 34: Räumliche Anordnung der Streudiagrammdaten.....	109
Abbildung 35: Clusterbewegung in TimeRider .....	110
Abbildung 36: Variablen sind nicht sortiert im Kontextmenü des Filters .....	111
Abbildung 37: prompting-Fehler in Usability.....	112
Abbildung 38: Schieberegler-Funktion der Zeitachse in TimeRider .....	113
Abbildung 39:Fading Effekt der Scatter Plots in TimeRider .....	114
Abbildung 40:Betroffene Funktionen der zusätzlichen Kategorie „Zeitachse“ .....	115
Abbildung 41: Häufigste Funktion der Kategorie „Abstract/Elaborate“ .....	116
Abbildung 42: Verwirrende Achsenbezeichnung in TimeRider.....	119
Abbildung 43: Unsortiertes Filtermenü in TimeRider .....	119
Abbildung 44:Überlagerung der Scatter Plots kann zur Verfärbung führen.....	120
Abbildung 45: overlapping führt zu Unübersichtlichkeit .....	121
Abbildung 46: Risikolevel Aktivierung, hebt den Normbereich hervor.....	122
Abbildung 47: präattentive ( <b>1&amp;2</b> ) & attentive Wahrnehmung ( <b>3</b> ) (in Anlehnung an Butz & Krüger, 2014, p.18). .....	155
Abbildung 48: Beispielbild für den Pop-out Effekt .....	156
Abbildung 49: Beispiel einer Small Multiples Visualisierung.....	157

## Tabellenverzeichnis

Tabelle 1: Auswertung der vierten Aufgabe (Lerneffekt).....	101
Tabelle 2: Auswertung der dritten Aufgabe (Animation). .....	103
Tabelle 3: Auswertung der zweiten Aufgabe (mapping).....	104
Tabelle 4: Eine Übersicht über alle kategorisierten Ereignisse von allen ProbandInnen..	104
Tabelle 5: Die überarbeitete Übersicht über alle kategorisierten Fehler nach Forsell & Johansson (2010) von allen ProbandInnen .....	105
Tabelle 6: Eine Übersicht über aller Programmfehler/ Abstürze welche je nach Testperson auftraten .....	106
Tabelle 7: Eine Übersicht über aller Programmabstürze welche je nach Testperson auftraten .....	106
Tabelle 8: Ranking der gemessenen Usability Fehler nach Forsell & Johansson, 2010...	108
Tabelle 9: Auswertung der Interaktionen nach Yi, et al. (2007) .....	115



## Literaturverzeichnis

Aigner, W., Miksch, S., Müller, W., Schumann, H. & Tominski, C., 2007. Visualizing time oriented data - a systematic view. *Computers & Graphics*, 31(3), pp.401-409.

Bakhtin, M.M., n.d. The problem of speech genres. In: C., Emerson & M., Holmquist, eds. 1986. *Speech Genres & Other Late Essays*. Translated by V. W. McGee. Austin(US-TX): Univ. Texas Press, pp.60–102.

Barth, A.P., 2013. *Algorithmik für Einsteiger: Für Studierende, Lehrer und Schüler in den Fächern Mathematik und Informatik*. 2nd ed. Wiesbaden (DE-HE): Springer Spektrum.

Bederson, B. & Shneiderman, B., 2004. *The Craft of Information Visualization: Readings and Reflections*. 2<sup>nd</sup> ed. San Francisco (US-CA): Elsevier.

Boren, T.M. & Ramey, J., 2000. Thinking Aloud: Reconciling theory and practice. *IEEE transactions on professional communication*, September 2000, 43 (3), pp.261-278.

Butz, A. & Krüger, A., 2014. *Mensch-Maschine-Interaktion*. München (DE-BY): De Gruyter - Oldenbourg.

Card, S.K., Mackinlay, J.D. & Shneiderman, B., 1999. *Readings in information visualization: using vision to think*. San Francisco (US-CA): Morgan Kaufmann Publishers Inc.

Cingi, C.C., 2013. Computer animation in teaching surgical procedures. In: *Proceedings of the 13th International Educational Technology Conference*. pp.230-237.

Cook, K.A., Earnshaw, R.A. & Stasko, J., 2007. Discovering the Unexpected. *Special Issue of IEEE Computer Graphics and Applications*, (27) 5, pp.15-18.

Ericsson, K.A. & Simon, H.A., 1980. Verbal reports as data. *Psychological Review*, May, 87(3), pp.215-251.

Ericsson, K.A. & Simon, H.A., 1993. *Protocol analysis: verbal reports as data*. 2<sup>nd</sup> ed. Boston (US-MA): MIT Press.

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., 1996. *From Data Mining to Knowledge Discovery in Databases*. AI Magazine, 17(3), pp.37-54.

Fels, U., 2015. *Usability-Analyse des Programms Animated Scatter Plot*. Mag. rer. soc. oec. Technische Universität Wien.

Forsell, C. & Johansson, J., 2010. An heuristic set for evaluation in information visualization. *In: Proceedings of the International Conference on Advanced Visual Interfaces*, pp.199-206.

Friel, S.N., Curcio, F.R. & Bright, G.W., 2001. Making sense of graphs: critical factors influencing comprehension and instructional implication. *Journal for Research in Mathematics Education*, March 2001, 32(2), pp.124-158.

Gershon, N., Eick, S.G. & Card, S., 1998. Information visualization: interactions. *ACM*, March 1998, 5(2), pp.9-15.

Gotz, D. & Zhou, M.X., 2008. Characterizing Users' Visual Analytic Activity for Insight Provenance. *IEEE Symposium on Visual Analytics Science and Technology*, 19-24 October 2008, pp.123-130.

Grimes, D.A. & Schulz, K.F., 2002. Cohort studies: marching towards outcomes. *The Lancet*, 26 January, 359(9303), pp.341-345.

Homepage der Anglia Ruskin Universität, 2008. *Guide to the Harvard Style of Referencing* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://libweb.anglia.ac.uk/referencing/harvard.htm>

[letzte Einsicht 02.10.2008]

Homepage der Firma Systema, 2013. *Startseite der Firma Systema* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.systema.info/>

[letzte Einsicht 2.05.2013]

Homepage der „HCI“, 2013. *Forschungsgruppe Human Computer Interaction des Instituts für Gestaltungs- und Wirkungsforschung, Technische Universität Wien* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://igw.tuwien.ac.at/hci/>

[letzte Einsicht 02.05.2013]

Homepage der „IFS“, 2006. *Homepage des IFS – Übungsaufgaben WS06/07* [Online] (letztes Update, 09.11.2006)

Erreichbar unter:

[http://ieg.ifs.tuwien.ac.at/~aigner/teaching/ws06/infovis\\_ue/infovis\\_ue\\_aufgabe3-techniken.html](http://ieg.ifs.tuwien.ac.at/~aigner/teaching/ws06/infovis_ue/infovis_ue_aufgabe3-techniken.html)

[letzte Einsicht 09.09.2015]

Homepage der „InfoVis-Wiki Seite“, 2006. *InfoVis-Wiki Seite mit einem Eintrag zu „Details on demand“* [Online] (letztes Update, 06.12.2006)

Erreichbar unter: [http://www.infovis-wiki.net/index.php?title=Details\\_on\\_demand](http://www.infovis-wiki.net/index.php?title=Details_on_demand)

[letzte Einsicht 22.09.2015]

Homepage der „LKNOE“, 2013. *Startseite der Landeskliniken Holding* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.holding.lknoe.at/>

[letzte Einsicht 02.05.2013]

Homepage der „NN/g“, 1995. *Startseite der Nielsen Norman Group - Evidence-Based User Experience Research, Training, and Consulting. 10 Usability Heuristics for User Interface Design by Jakob Nielsen*, [Online] (letztes Update, 01.01.1995)

Erreichbar unter: <http://www.nngroup.com/articles/ten-usability-heuristics/>

[letzte Einsicht 06.02.2013]

Homepage der „NN/g“, 2000. *Seite der Nielsen Norman Group - Why You Only Need to Test with 5 Users by Jakob Nielsen*, [Online] (letztes Update, 19.03.2000)

Erreichbar unter: <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>

[letzte Einsicht 21.08.2015]

Homepage der „NN/g“, 2014. *Startseite der Nielsen Norman Group - Evidence-Based User Experience Research, Training, and Consulting. Memory Recognition and Recall in User Interfaces by Raluca Budiu*, [Online] (letztes Update, 06.07.2014)

Erreichbar unter: <http://www.nngroup.com/articles/recognition-and-recall/>

[letzte Einsicht 24.01.2015]

Homepage der „NN/g“, 2015. *Startseite der Nielsen Norman Group - Evidence-Based User Experience Research, Training, and Consulting. Thinking Aloud: The #1 Usability Tool by Jakob Nielsen*, [Online] (letztes Update, 16.01.2012)

Erreichbar unter: <http://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>

[letzte Einsicht 29.05.2015]

Homepage der "The history of a picture's worth", 2015. *Printers' Ink, eine Anzeige mit dem Slogan „One Picture is Worth A Thousand Words* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www2.cs.uregina.ca/~hepting/research/web/words/history.html>

[letzte Einsicht 07.08.2015]

Homepage der Universität des Saarlandes, 2010. *Informationswissenschaft (FR 5.6)/Themen/Daten, Wissen, Information/Definition: Information* [Online] (letztes Update, 14.08.2014)

Erreichbar unter: <http://www.uni-saarland.de/campus/faculties/departments/humanities-iii/departments/informationswissenschaft/infowissthemem/datwissinf/definitioninformation.html?type=99%3Ftype%3D99>

[letzte Einsicht 14.08.2014]

Homepage der Universität Duisburg-Essen, 2015. *Einführung: Behaviorismus* [Online] (letztes Update, n.p.)

Erreichbar unter: <https://www.uni-due.de/edit/lp/behavior/behavior.htm>

[letzte Einsicht 25.08.2015]

Homepage der Universität Regensburg, 2015. *Einführung in Usability* [Online] (letztes Update, n.p.)

Erreichbar unter: [http://www.uni-regensburg.de/Fakultaeten/phil\\_Fak\\_II/Psychologie/Doktoranden/absolventen/eichinger\\_armin/usability.html](http://www.uni-regensburg.de/Fakultaeten/phil_Fak_II/Psychologie/Doktoranden/absolventen/eichinger_armin/usability.html)

[letzte Einsicht 11.02.2015]

Homepage der Universität Rostock, 2002. *Usability - Usability Laboratorien* [Online] (letztes Update, n.p.)

Erreichbar unter: [http://wwwswt.informatik.uni-rostock.de/deutsch/Mitarbeiter/michael/lehre/Usab\\_WS2002/Jan/vortrag\\_html.htm](http://wwwswt.informatik.uni-rostock.de/deutsch/Mitarbeiter/michael/lehre/Usab_WS2002/Jan/vortrag_html.htm)

[letzte Einsicht 11.02.2015]

Homepage der Universität zu Lübeck, 2014. *Homepage des Projekts: 4D Medical Image Computing zur modellbasierten Analyse atmungsbedingter Tumor- und Organbewegungen* [Online] (letztes Update, 26.06.2014)

Erreichbar unter: <https://www.imi.uni-luebeck.de/de/content/4d-medical-image-computing-zur-modellbasierten-analyse-atmungsbedingter-tumor-und-organbeweg>

[letzte Einsicht 26.09.2015]

Homepage des "ACM SIGGRAPH Education Committee", 2007. *Offizielle Webseite der Association for Computing Machinery und deren Special Interest Group for Computer Graphics and Interactive Techniques Bildungsausschusses zum Thema Definitionen und Grundlagen der Visualisierungen* [Online] (letztes Update, 11.02.1999)

Erreichbar unter:

<http://www.siggraph.org/education/materials/HyperVis/visgoals/visgoal2.htm>

[letzte Einsicht 11.09.2010]

Homepage des „CVASt“, 2013. *Projektseite des Centre for Visual Analytics Science and Technology der Information Engineering Group (IEG) des Institute of Software Technology and Interactive Systems (IFS) der Technische Universität Wien* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.cvast.tuwien.ac.at/cvast>

[letzte Einsicht 01.03.2013]

Homepage des deutschen Wörterbuchverlags „PONS“, 2013. *Online Wörterbuchsuche für das Wort „data mining“* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://de.pons.com/übersetzung?q=data+mining&l=deen&in=&lf=de>

[letzte Einsicht 08.07.2013]

Homepage des „DIKE“, 2011. *Startseite des Department für Information und Knowledge Engineering der Donau-Universität Krems.* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.donau-uni.ac.at/de/department/ike/index.php>

[letzte Einsicht 16.10.2011]

Homepage des „Duden“, 2013a. *Online Wörterbuchsuche für das Wort „multivariat“* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/multivariat>

[letzte Einsicht 28.11.2013]

Homepage des „Duden“, 2013b. *Online Wörterbuchsuche für das Wort „animieren“* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/animieren>

[letzte Einsicht 12.01.2013]

Homepage des „Duden“, 2013c. *Online Wörterbuchsuche für das Wort „Schwarm“* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Schwarm>

[letzte Einsicht 12.01.2013]

Homepage des „Duden“, 2013d. *Online Wörterbuchsuche für das Wort „Regression“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Regression>

[letzte Einsicht 12.01.2013]

Homepage des „Duden“, 2013e. *Online Wörterbuchsuche für das Wort „Korrelation“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Korrelation>

[letzte Einsicht 12.01.2013]

Homepage des „Duden“, 2015a. *Online Wörterbuchsuche für das Wort „Rendering“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Rendering>

[letzte Einsicht 13.08.2015]

Homepage des „Duden“, 2015b. *Online Wörterbuchsuche für das Wort „Zeitstrahl“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Zeitstrahl>

[letzte Einsicht 12.08.2015]

Homepage des „Duden“, 2015c. *Online Wörterbuchsuche für das Wort „Mapping“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Mapping>

[letzte Einsicht 13.08.2015]

Homepage des „Duden“, 2015d. *Online Wörterbuchsuche für das Wort „Behaviorismus“*  
[Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.duden.de/rechtschreibung/Behaviorismus>

[letzte Einsicht 25.08.2015]

Homepage des Fraunhofer- IGD, 2013. *Fraunhofer - Institut für Graphische Datenverarbeitung - Abteilung Informationsvisualisierung und Visual Analytics* [Online] (letztes Update, n.d.)

Erreichbar unter: <https://www.igd.fraunhofer.de/Institut/Abteilungen/IVA>

[letzte Einsicht 02.05.2013]

Homepage des freien online Lexikons „Webopedia“, 2013. *Online Lexikon Eintrag zu „Logfile“* [Online] (letztes Update, n.d.)

Erreichbar unter: [http://www.webopedia.com/TERM/L/log\\_file.html](http://www.webopedia.com/TERM/L/log_file.html)

[letzte Einsicht 29.11.2013]

Homepage des „Handbuch Usability“, 2007. *Online Leitfaden zum Thema Usability* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.handbuch-usability.de/iso-9241.html>

[letzte Einsicht 03.05.2013]

Homepage des internationalen Computerbuchverlages O'Reilly Media, 2012. *What is big data? An introduction to the big data landscape*. By Dumbill, E., O'Reilly Radar [Online] (letztes Update, n.d.)

Erreichbar unter: <http://radar.oreilly.com/2012/01/what-is-big-data.html>

[letzte Einsicht 26.05.2013]

Homepage des „KSA“, 2013. *Startseite des Instituts für Kultur- und Sozialanthropologie der Universität Wien - Grundlagen statistischer Auswertungsverfahren von Erwin Ebermann* [Online] (letztes Update, 18.06.2010, 16:42)

Erreichbar unter: <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-91.html>

[letzte Einsicht 04.02.2013]

Homepage des „LKNOE-UK Krems“, 2013. *Startseite des Universitätsklinikum Krems* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.krems.lknoe.at/>

[letzte Einsicht 02.05.2013]

Homepage des „NÖGUS“, 2013. *Startseite des niederösterreichischen Gesundheits- und Sozialfonds* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.noegus.at/>

[letzte Einsicht 02.05.2013]

Homepage des Oxford English Dictionary Online, 2015. *Online Wörterbuchsuche für das Wort „black box“ - Oxford University Press*. [Online] (letztes Update, Juni 2015)

Erreichbar unter: <http://www.oed.com/view/Entry/282116?redirectedFrom=black+box#eid>

[letzte Einsicht 25 August 2015]

Homepage des Programms „TimeRider“, 2011. *Projektseite der Information Engineering Group (IEG) des Institute of Software Technology and Interactive Systems (IFS) der TU Wien* [Online] (letztes Update, 22.07.2011)

Erreichbar unter: <http://ieg.ifs.tuwien.ac.at/projects/VisualizingPatientCohorts/>

[letzte Einsicht 09.08.2011]

Homepage des Projekts „VisuExplore“, 2012. *Projektseite der Information Engineering Group (IEG) des Institute of Software Technology and Interactive Systems (IFS) der TU Wien* [Online] (letztes Update, 03.02.2012)

Erreichbar unter: <http://ieg.ifs.tuwien.ac.at/projects/VisuExplore/>  
[letzte Einsicht 06.05.2012]

Homepage des „Xerox PARC“, 2013. *Startseite des Xerox Palo Alto Research Center* [Online] (letztes Update, n.d.)

Erreichbar unter: <http://www.parc.com/>  
[letzte Einsicht 06.06.2013]

Homepage von Edward Tufte, 2015. *Edward Tufte forum: Graphical summaries for medical patients* [Online] (letztes Update, n.d.)

Erreichbar unter: [http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0003mm](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0003mm)  
[letzte Einsicht 09.09.2015]

Homepage von „InfoVis.net“, 2015. *Homepage des „The Golden Age of Visualisation“ by Dürsteler, J.C.* [Online] (letztes Update, 09.01.2003)

Erreichbar unter: <http://www.infovis.net/printMag.php?num=111&lang=2>  
[letzte Einsicht 14.08.2015]

Homepage von „nowscape.com“, 2015. *Homepage mit der Darstellung des Napoleon-Russland-Feldzugs* [Online] (letztes Update, 27.01.2014)

Erreichbar unter: [http://nowscape.com/star\\_city/map\\_napoleon\\_moscow.html](http://nowscape.com/star_city/map_napoleon_moscow.html)  
[letzte Einsicht 13.08.2015]

Keim, D.A., Mueller, W., & Schumann, H., 1997. *Visual data mining*. Konstanz (DE-BW): Bibliothek der Universität Konstanz.

Keim, D.A., North, S., Panse, C. & Sips, M., 2004. Visual data mining in large geo-spatial point sets. *IEEE Computer Graphics and Application*, 12, pp.36–44.

Krahmer, E. & Ummelen, N., 2004. Thinking about thinking aloud: a comparison of two verbal protocols for usability testing. *IEEE Transactions on Professional Communication*, 47 (2), pp.105-117.

Kriglstein, S., Pohl, M. & Stachl, C., 2012. Animation for Time-oriented Data: An Overview of Empirical Research. *16th International Conference on Information Visualisation (IV)*. Montpellier (FR-K), 11-13 July 2012, IEEE, pp.30-35.

- Kriglstein, S., Pohl, M. & Smuc, M., 2014. Pep Up Your Time Machine: Recommendations for the Design of Information Visualizations of Time-Dependent Data. In: W. Huang, ed. 2014. *Handbook of Human Centric Visualization*. Springer New York, pp.203-225.
- Lewis, C.H., 1982. *Using the "Thinking Aloud" method in cognitive interface design*. Yorktown Heights, Westchester (US-NY): IBM T.J. Watson Research Center.
- Marey, E.J., 1885. *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine*. Paris (FR-J): G. Masson.
- Martins, S.B., Shahar, Y., Goren-Bar, D., Galperin, M., Kaizer, H., Basso, L.V., McNaughton, D. & Goldstein, M.K., 2008. Evaluation of an architecture for intelligent query and exploration of time-oriented clinical data. *Artificial Intelligence in Medicine*, 43(1), May 2008, pp.17-34.
- Mazza, R., 2009. *Introduction to Information Visualization*. London (GB-LND): Springer-Verlag London Limited.
- McCormick, B.H., DeFanti, T.A., Brown, M.D., 1987. Visualization in Scientific Computing. *ACM Siggraph, Computer Graphics*, 21(6), pp.1-14.
- Miller, G.A., 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, 63(2), pp.81–97.
- Neubauer, B., 2012. *A Comparison of Static and Dynamic Visualizations for Time-Oriented Data*. Diplom-Ingenieurin. Technische Universität Wien.
- Nielsen, J., & Landauer, T.K., 1993. A mathematical model of the finding of usability problems. In: *Proceedings of the ACM/IFIP INTERCHI'93 Conference*. 24-29 April, Amsterdam(NL-NH), pp.206-213.
- Nielsen, J., 1994a. Enhancing the explanatory power of usability heuristics. In: *Proceedings ACM CHI'94 Conf*. Boston (US-MA), 24-28 April, pp.152-158.
- Nielsen, J., 1994b. Heuristic evaluation. In J. Nielsen and R. L. Mack (eds.). *Usability inspection methods*. Usability Inspection Methods. NewYork (US-NY): John Wiley & Sons. pp.25-61.
- Nielsen, J., 1994c. *Usability Engineering*. Morgan Kaufmann.
- Nielsen, J., 1999. *Designing Web Usability*. Peachpit Press. Safari Tech Books Online.

- Nielsen, J., & Molich, R., 1990. Heuristic evaluation of user interfaces, *In: Proceedings ACM CHI'90 Conf.* Seattle (US-WA), 1-5 April, pp.249-256.
- Nischwitz, A., Fischer, M., Haberbaecker, P. & Socher, G., 2011. *Computergrafik und Bildverarbeitung: Band I: Computergrafik*. 3rd ed. Wiesbaden (DE-HE): Vieweg+Teubner Verlag, Springer Fachmedien Wiesbaden GmbH.
- Plaisant, C., Mushlin, R., Snyder, A., Lia, J., Heller D. & Shneiderman B., 1998. *LifeLines: Using Visualization to Enhance Navigation and Analysis of Patient Records*. October 1998, HCIL Technical Report No. 98-08
- Preim, B. & Bartz, D., 2007. *Visualization in medicine: theory, algorithms, and applications. The Morgan Kaufmann Series in Computer Graphics*. San Francisco (US-CA): Morgan Kaufmann Publishers Inc, Elsevier Inc.
- Pike, A.W.G., Hoffmann, D.L., García-Diez, M., Pettitt, P.B., Alcolea, J., De Balbín, R., González-Sainz, C., de las Heras, C., Lasheras, J.A., Montes, R. & Zilhão, J., 2012. U-Series Dating of Paleolithic Art in 11 Caves in Spain: Materials/Methods, Supporting Text, Tables, Figures, and/or References. *Science*, 15 June, 336(6087), pp.1409-1413.
- Pohl, M., Wiltner, S. & Miksch, S., 2010. Exploring information visualization: describing different interaction patterns. *In: Proceedings of the 3rd Workshop: Beyond time and errors: novel evaluation methods for Information Visualization (BELIV '10)*, ACM Press, pp.16-23.
- Pohl, M., Wiltner, S., Rind, A., Aigner, W., Miksch, S., Turic, T., & Drexler, F., 2011. Patient development at a glance: an evaluation of a medical data visualization. *In: Proceedings of 13th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2011)*, Part IV, Heidelberg (DE-BW): Springer, pp.292-299.
- Pohl, M., Wiltner, S., Miksch, S., Aigner, W., Rind, A., 2012. Analysing interactivity in information visualisation. *Zeitschrift KI - Künstliche Intelligenz*, 26(2), pp.151-159.
- Powsner, S.M. & Tufte, E.R., 1994. Graphical summary of patient status. *The Lancet*, 6 Aug, 344(8918), pp.386-389.
- Ratwani, R.M., Trafton, J.G. & Boehm-Davis, D.A., 2008. Thinking graphically: connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology*, March, 14(1), pp.36-49.

- Rind, A., Aigner, W., Miksch, S., Wiltner, S., Pohl, M., Drexler, F., Neubauer, B. & Suchy, N., 2011. Visually Exploring Multivariate Trends in Patient Cohorts using Animated Scatter Plots. *M.M. Robertson (ed.), Ergonomics and Health Aspects of Work with Computers, In: Proceedings of the International Conference held as part of HCI International 2011*, LNCS 6779, pp.139-148, Heidelberg (DE-BW): Springer. doi:10.1007/978-3-642-21716-6\_15.
- Robertson, G., Fernandez, R., Fisher, D., Lee, B. & Stasko, J., 2008. Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), pp.1325-1332.
- Ruiz, J.G., Cook, D.A. & Levinson, A.J., 2009. Computer animations in medical education: a Critical literature review. *Medical Education*, 43(9), pp.838-846.
- Schumann, H., 2004. Konzepte und Methoden der wissenschaftlichen Visualisierung. *Kartographische Bausteine - TU Dresden*, 26, pp.20-28.
- Schütz, T., 2013. *Hucho - Aerodynamik des Automobils : Strömungsmechanik, Wärmetechnik, Fahrdynamik, Komfort*. 6<sup>th</sup> ed. Wiesbaden (DE-HE): Springer Fachmedien Wiesbaden.
- Senay, H. & Ignatius, E., 1994. A knowledge-based system for visualization design. *Computer Graphics and Applications, IEEE*, 14(6), pp.36-47.
- Shackel, B. & Richardson, S.J., 1991. *Human Factors for Informatics Usability*, New York (US-NY): Cambridge University Press.
- Shneiderman, B. & Plaisant, C., 2009. *Designing the user interface: Strategies for effective Human-Computer Interaction*. 5<sup>th</sup> ed. Boston (US-MA): Addison-Wesley Publishing Co., Inc is an imprint of Pearson.
- Spence, R., 2007. *Information Visualization: Design for Interaction*. 2nd ed. Upper Saddle River (US-NJ): Prentice-Hall, Inc.
- Suchy, N., 2010. *Eine Bachelorarbeit über die Thinking Aloud Methode, im Rahmen eines Usability Testings unter Verwendung von Human-Computer Interaction Techniken und Erkenntnissen*. B.Sc. Technische Universität Wien.
- Trafton, J.G., Kirschenbaum, S.S., Tsui, T.L., Miyamoto, R.T., Ballas, J.A. & Raymond, P.D., 2000. Turning pictures into numbers: extracting and generating information from complex visualizations. *International Journal of Human Computer Studies*, 53(5), pp.827-850.

Trickett, S.B. & Trafton, J.G., 2006. Toward a comprehensive model of graph comprehension: Making the case for spatial cognition. *Diagrammatic Representation and Inference: 4th International Conference, Diagrams 2006, Stanford (US-CA), USA*, 4045, pp.286-300.

Unger, J., 2014. *Visual analytics of clinical data and treatment processes for cohorts*. Diplom-Ingenieur. Technische Universität Wien.

Valdes, V.C. & Bischoff, J.L., 1989. Accelerator 14C dates for early upper paleolithic (basal Aurignacian) at El Castillo Cave (Spain). *Journal of Archaeological Science*, 16, pp.577-584.

Watson, J.B., 1913. Psychology as the Behaviorist Views it. *Psychological Review*, March, 20(2), pp.158-177.

Weiß, C., 2013. *Basiswissen Medizinische Statistik*. Heidelberg (DE-BW): Springer Berlin Heidelberg.

Wong, P.C. & Thomas, J., 2004. Visual Analytics. *IEEE Computer Society: IEEE Computer Graphics and Applications*, 24 (5), pp.20–21.

Yi, J.S., Kang, Y.A., Stasko, J.T. & Jacko, J.A., 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on visualization and computer graphics*, 13(6), pp.1224-1231.

## A Anhang

Der Anhang A enthält die Dokumente, welche vor und während der Usability Untersuchung von den ProbandInnen verwendet wurden. Das erste Dokument „TimeRider Einschulung“ war eine kurze schriftliche Einführung in die Untersuchung und wurde von den Testpersonen vor dem Test durchgelesen, um einen Überblick zu gewinnen. Das zweite Dokument „Parameterdefinition“, beschrieb die humanmedizinischen Werte, welche im Test vorkommen konnten. Das dritte Dokument war eine „Liste der Variablen“, welche alle im Programm verwendeten medizinischen Abkürzungen auflistete. Das letzte Dokument bilden „Die vier Aufgaben“, welche von den ÄrztInnen im Zuge des Tests gelöst werden sollten.



## A.1 TimeRider Einschulung

### 1. Worum geht es?

Zu allererst vielen Dank für Ihre freiwillige Teilnahme an der Versuchsreihe mit dem neuesten Programm auf dem Gebiet der medizinischen Informationsvisualisierung. Das vorliegende Programm visualisiert Parameter von Patientendaten, die aufgrund eines diagnostizierten Diabetes Mellitus (Typ I/II) in medizinischer Behandlung stehen und teilweise ein bestimmtes Medikament zur Einnahme verschrieben bekommen haben.

Bei jedem Kontrolltermin, den die Patienten wahrnehmen, werden (idealerweise) die Parameter festgehalten, welche auf einem der beiden zusätzlichen Informationsblätter kurz erklärt werden.

Die Kontrolltermine und die Abstände zwischen den Terminen variieren. Daher gibt es unterschiedlich viele Messungen der Parameter bei verschiedenen Patienten. Ebenso kam es immer wieder zu Unterbrechungen der Messungen, durch die natürliche Fluktuation des Patientenstromes aufgrund von vorzeitigem Therapieabbruch, Genesung sowie Exitus seitens der Patienten.

Das vorliegende Programm soll eine Tendenz (einen Trend) erkennbar machen, welche sich durch die Behandlung (z.B. Medikation) der Diabetes-Patienten mit der Zeit abzeichnet (sofern die Behandlung eine Auswirkung auf die erhobenen Parameter hat).

Diese Tendenz wird mit der Anzahl der Patienten und mit der Anzahl der Kontrolltermine pro Patient naturgemäß deutlicher.

### 2. Wie funktioniert das Programm?

Zusammenhang:

In einem Koordinatensystem werden jeweils zwei Parameter miteinander in Verbindung gebracht, zum Beispiel Körpergewicht (x-Achse) und Cholesterin (y-Achse).

Animation:

Diese Darstellung kann nun animiert werden (ähnlich einem Video, das abgespielt wird). Start ist der Zeitpunkt der ersten Kontrolluntersuchung des ersten Patienten. Ende ist der Zeitpunkt der aktuell letzten Kontrolluntersuchung des aktuell letzten Patienten.

Tendenz, Trend:

Die erfassten Messwerte werden zum Zeitpunkt ihrer Erfassung als Punkt im Koordinatensystem eingeblendet. Es handelt sich also nicht um eine statische Darstellung des Zusammenhanges zweier Parameter, sondern um eine sich mit der Zeit wandelnde Darstellung, die eventuelle Änderungen (Tendenzen, Trends) erkennbar machen soll.

Das Programm ist ein Prototyp. Daher kann es zu Fehlern und Fehldarstellungen kommen. Bitte, dies zu berücksichtigen.

### 3. Was sollen Sie tun?

Wir ersuchen Sie das vorliegende Programm auszuprobieren.

Zuerst kann das Programm „frei“ ausprobiert werden um ein Gefühl für das Handling der Software zu entwickeln. Nach dieser kurzen Probephase werden ihnen vier Aufgabenstellungen vorgelegt, welche sie mittels des neuen Programms lösen sollen. Sollte sich eine Aufgabe als unlösbar erweisen, wird eine Hilfestellung geboten oder die Aufgabe abgebrochen.

Die verwendete Methode zur Identifizierung von Schwachstellen in der Benutzbarkeit des Programms wird als „thinking aloud“ betitelt, zu Deutsch: „Lautes Denken“. Dabei sollen die ProbandInnen möglichst alle Gedankengänge laut aussprechen und seine Aktionen und Probleme artikulieren und dadurch Probleme im Bereich der Benutzbarkeit aufzeigen. Dabei ist es wichtig, dass im Speziellen die Gedanken in Phasen der Orientierung und die negativen Eindrücke wiedergegeben werden, z.B. wenn man Elemente der Benutzeroberfläche nicht findet weil sie am „falschen“ Ort sind.

Es ist wichtig, sich im Klaren zu sein, dass nur die Benutzbarkeit des Programms getestet wird und auf keinen Fall über die Fähigkeit der Testpersonen geurteilt wird. Ebenfalls ist negative Kritik durchaus angebracht und wird nicht als persönliche Kritik verstanden. Finden SIE sich nicht zurecht, ist das nicht Ihr Fehler, sondern eine Schwachstelle im Programm.

Da während des Tests nur ein Monolog stattfindet, wird nach dem Abschluss aller Aufgaben ein kurzes Interview stattfinden, in dem Erfahrungen und Meinungen kundgetan werden können.

Am Ende möchten wir uns nochmals für Ihre Mitarbeit und ihre Zeit bedanken, die Sie für dieses Projekt bereit sind aufzubringen.

## A.2 Parameterdefinition

**Größe [cm]** – Körpergröße, welche hier in Zentimetern angegeben wird

**BMI - Body Mass Index:** Maß für das Gesundheitsrisiko durch das Körpergewicht.

*Rechnungsbeispiel: Größe: 1,70 m, Gewicht: 60 kg = 60 geteilt durch 1,70 = 20,76*

Werte laut **WHO** (Weltgesundheitsorganisation):

	Untergewicht	Normalgewicht	Leichtes Übergewicht	Erhebliches Übergewicht
Weiblich	Unter 19	19-24	24-29	Über 29
Männlich	Unter 20	20-25	25-30	Über 30

**HbA1c [%]** - Glykolisiertes Hämoglobin in Ketoaminform: Normalbereich zwischen 29 und 42 mmol/mol

**RR syst [mmHG]** - **Riva Rocci systolisch** (kontrahiert): Arterieller Druck, der durch die Auswurfkraft des Herzens erzeugt wird Normalbereich Ruhelage: 100-130 mmHg

**RR diast [mmHG]** - **Riva Rocci diastolisch** (erschläft): Minimaler arterieller Druck während der Herzfüllungsphase Normalbereich Ruhelage: 60-85 mmHg

**Crea** - Creatinin (Nierenwert): Normale Ausscheidung im Harn: 1,0 – 1,5 g pro 24h

**Chol** – Cholesterinwert: Cholesterinspiegel, Durchschnittswert zw. 35 und 65 Jahren: 236 mg/dl Bei jüngeren Patienten ist der Cholesterinwert niedriger.

**TG** - Triacylglycerine (Blutfettwerte): Erhöhte Werte über 150 mg/dl weisen auf eine Fettstoffwechselstörung, Diabetes oder Übergewicht hin.

**Gewicht** - Körpergewicht in Kilogramm

**NBZ [mg/dl]** – Nüchternblutzucker: Risikobereich >100 mg/dl. Bei Patienten mit Diabetes Mellitus liegt der NBZ Wert normalerweise im Risikobereich.

**Alter** - In Jahren

**Diabetes Typ 1** betrifft normalerweise Kinder und Jugendliche

**Diabetes Typ 2** Betrifft normalerweise Personen ab ~40 Jahren (Diese Grenzen verschwimmen aber zunehmend).



### A.3 Liste der Variablen

	Größe		biometrische Daten
	Gewicht		biometrische Daten
	BMI		biometrische Daten
NBZ	Blutzucker nüchtern	Zucker	Labor
HbA1c	Glykohämoglobin	Zucker	Labor
Chol	Cholesterin	Lipid	Labor
TG	Triglycerid	Lipid	Labor
Crea	Creatinin	Nierenfunktion	Labor
Harn	Eiweiß im Harn	Nierenfunktion	Labor
RR syst	Blutdruck: systolisch	Blutdruck	biometrische Daten
RR diast	Blutdruck: diastolisch	Blutdruck	biometrische Daten
	Nikotin		Risikofaktor
	Retinopathie	Augenschaden im Augenhintergrund	Endorganschaden
	Insult	Schlaganfall	Endorganschaden
KHK	Koronare Herzkrankheit		Endorganschaden
PAVK	Periphere Arterielle Verschlusskrankheit	Durchblutungsstörung	Endorganschaden
	Fußkontrolle	offene Stellen, Druckgeschwüre	Endorganschaden
PNP	Polyneuropathie	Erkrankungen des Nervensystems	Endorganschaden
	Nephropathie	Erkrankungen der Niere	Endorganschaden
SH	SH	Zucker	Therapie
Met	Met	Zucker	Therapie
Glit	Glit	Zucker	Therapie
DPP4	DPP4	Zucker	Therapie
Acarb	Acarb	Zucker	Therapie
VZI	VZI	Insulin	Therapie
ALT	ALT	Insulin	Therapie
Misch	Misch	Insulin	Therapie
ACE	ACE	Blutdruck	Therapie
Bbl	Bbl	Blutdruck	Therapie
RR sonst	RR sonst	Blutdruck	Therapie
Statin	Statin	Blutfette	Therapie
ASS	ASS	Durchblutung	Therapie



## A.4 Die vier Aufgaben

### Task I

Vorgabe:

x-Achse: BMI  
y-Achse: HbA1c [%]

Aufgabenstellung:

Experimentieren Sie nach Belieben mit dem Programm und machen Sie sich vertraut mit der Oberfläche und den Funktionen.  
Beschreiben Sie ihre Erkenntnisse.

### Task II

Vorgabe:

x-Achse: BMI  
y-Achse: RR syst [mmHG]

Aufgabenstellung:

Heben Sie das Geschlecht der Patienten hervor (z.B. mittels Farbe oder Form).  
Zum letzten Untersuchungszeitpunkt könnte man die Patienten in zwei Gruppen einteilen, abhängig von Ihrem Gesundheitszustand. Welche Patienten haben gute Werte?  
Wählen Sie jeweils eine typische (durchschnittliche) Person aus den zwei Punktwolken aus.  
Begründen Sie ihre Auswahl.

### Task III

Vorgabe:

x-Achse: NBZ  
y-Achse: RR diast [mmHG]

Aufgabenstellung:

Schränken Sie den angezeigten Datensatz wie folgt ein:  $NBZ \leq 100$ ;  $RR \text{ diast.} \leq 80$ .  
Wählen Sie eine Darstellung, die einen guten Überblick über den Verlauf aller verbleibenden Patienten ermöglicht.  
Bei welchen Personen ist eine günstige Entwicklung zu erkennen?  
Wie entwickelt sich der allgemeine Gruppentrend?  
Experimentieren Sie nach Belieben.  
Beschreiben Sie ihre Erkenntnisse.

### Task IV

Vorgabe:

Keine

Aufgabenstellung:

Nutzen Sie gewonnene Eindrücke und Erkenntnisse, um beliebig mit dem Programm zu experimentieren und gegebenenfalls Unklarheiten zu beseitigen oder Interessantes genauer auszuprobieren.  
Beschreiben Sie ihre Erkenntnisse.



## B Anhang

### *Exkurs: Attentive & präattentive Wahrnehmung*

Die untenstehende Information ist aus dem Buch „Mensch-Maschine-Interaktion“ von Butz & Krüger, 2014.

Bestimmte visuelle Informationen (z.B. Formen-, Farb- und Bewegungswahrnehmung) werden bereits direkt im Nervensystem des menschlichen Auges und noch vor dem visuellen Cortex im menschlichen Gehirn verarbeitet. Somit ist diese Art der Wahrnehmung hochgradig parallel und die Wahrnehmungsprozesse laufen einer konstanten Zeit von ca. 200-250ms ab (unabhängig von der Reizanzahl). Da dieser Wahrnehmungsvorgang stattfindet bevor er im menschlichen Gehirn und somit die Aufmerksamkeit des Menschen erreicht, werden diese Prozesse als „**präattentive Wahrnehmung**“ (aus lat. *prä*: vor und *attentio*: **Aufmerksamkeit**) bezeichnet. Jene Wahrnehmung welcher Aufmerksamkeit gewidmet wird, heißt somit „**attentive Wahrnehmung**“ (Butz & Krüger, 2014). Die folgenden Bilder 1&2 der Abbildung 47 zeigen hierfür ein Beispiel:

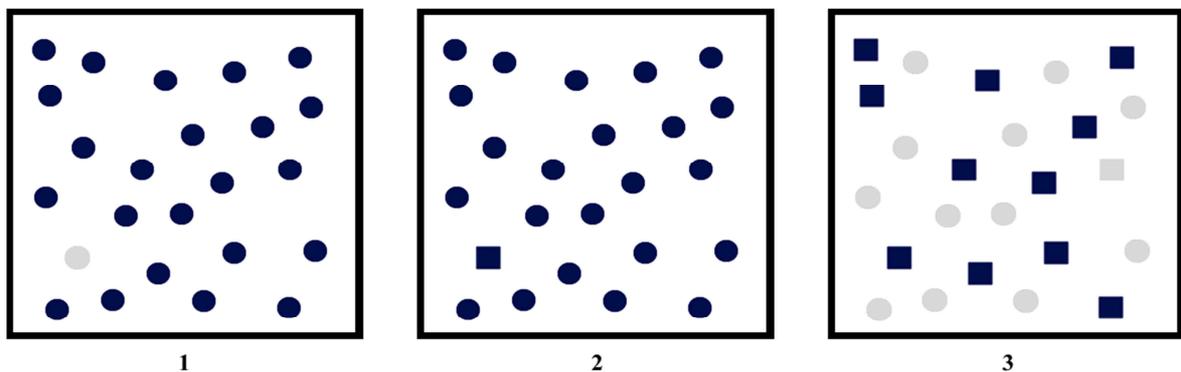


Abbildung 47: präattentive (1&2) & attentive Wahrnehmung (3) (in Anlehnung an Butz & Krüger, 2014, p.18).<sup>36</sup>

Der hellgraue Kreis in Bild 1 der Abbildung 47 und das dunkelblaue Quadrat in Bild 2 der Abbildung 47, springen den BetrachterInnen sofort ins Auge. Ohne die Bilder lange abzusuchen oder zu vergleichen wird sofort erkannt, wo sich das nicht zugehörige Objekt befindet. Die Wahrnehmung von Farbe und Form geschieht in einem präattentiven Prozess. Bei einer Suche nach einer Kombination von zwei oder mehreren Merkmalen, kann die präattentive Wahrnehmung nicht mehr verarbeiten. Das gesuchte Objekt in Bild 3 der Abbildung 47, ein hellgraues Rechteck, wird nicht sofort erkannt und der Betrachter muss daher alle Objekte sequentiell mit den beiden Merkmalen Helligkeit (hell vs. dunkel) und

<sup>36</sup> Beispielbilder für präattentive Wahrnehmung (1&2) und attentive Wahrnehmung (3). In Bild 3 ist es für einen Betrachter schwerer, das Objekt mit einer unterschiedlichen Form im Vergleich zu den anderen Objekten zu finden (Butz & Krüger, 2014, Mensch-Maschine-Interaktion, p.18).

Form (Quadrat) untersuchen, um es zu finden. Mit der Anzahl der dargestellten Objekte wächst die benötigte Zeit linear und beim Betrachten von Bild 3 der Abbildung 47, welches nur 24 Objekte enthält, dauert es für einen Betrachter einige Zeit bis das hellgraue Quadrat entdeckt wird. Des Weiteren sind Merkmale wie Größe, Orientierung, Krümmung, Bewegungsrichtung und räumliche Tiefe ebenfalls präattentiv wahrnehmbar. Daher ist bei der Verwendung von Objekten bei grafischer Gestaltung einer Präsentation zu beachten, dass wenn sie sich in genau einem präattentiv wahrnehmbaren Merkmal unterscheiden, sie sehr schnell selbst aus einer großen Menge von Objekten ersichtlich sind und bei einer Kombination von solchen Merkmalen nur nach genauen Suchen auffindbar sind. Somit können bei gezieltem Einsatz dieses Wissens, Informationen in einer grafischen Darstellung gezielt erkenntlich gemacht werden (Pop-out Effekt, siehe Abbildung 48) oder eben auch gut versteckt werden (siehe Bild 3 in Abbildung 47) (Butz & Krüger, 2014).

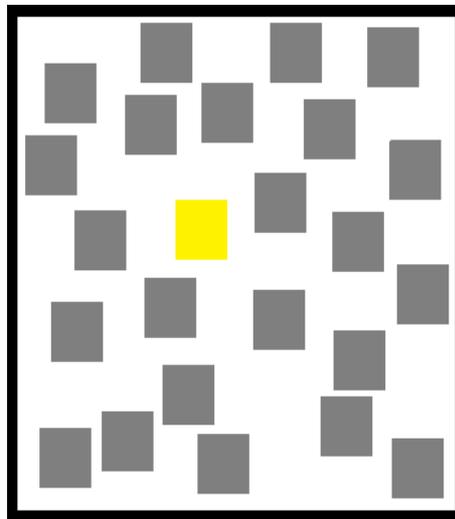


Abbildung 48: Beispielbild für den Pop-out Effekt<sup>37</sup>

---

<sup>37</sup> Um den Fokus auf ein bestimmtes Objekt zu legen nutzt man eine gezielt hervorgehobene Darstellung. Dem Betrachter springt das gelbe Quadrat sofort ins Auge (Quelle: Eigene Darstellung).

## C Anhang

### *Exkurs: Small Multiples*

„*Small Multiples*“ ist ein verbreitetes Visualisierungskonzept, wo mehrere kleine Darstellungen direkt nebeneinander angeordnet werden (siehe Abbildung 49). Jede der kleinen Repräsentationen zeigt dieselbe Konfiguration (z.B. selber Wertebereich auf den Achsen), aber zu unterschiedlichen Zeitpunkten oder Objekten. Tufte (Tufte, 1983 zitiert in Neubauer, 2012) stellte in seiner Arbeit das sogenannte „*Small Multiples*“ Konzept vor, welches auf dem physiologischen Prinzip der menschlichen Wahrnehmung basiert. Tufte argumentierte, dass das menschliche Auge sehr gut Unterscheidungen auf kleinsten Raum durchführen kann, z.B. können nahe aneinander liegende Linien oder 100 einfarbige Punkte auf einem Quadratzentimeter einfach unterschieden werden (Neubauer, 2012).

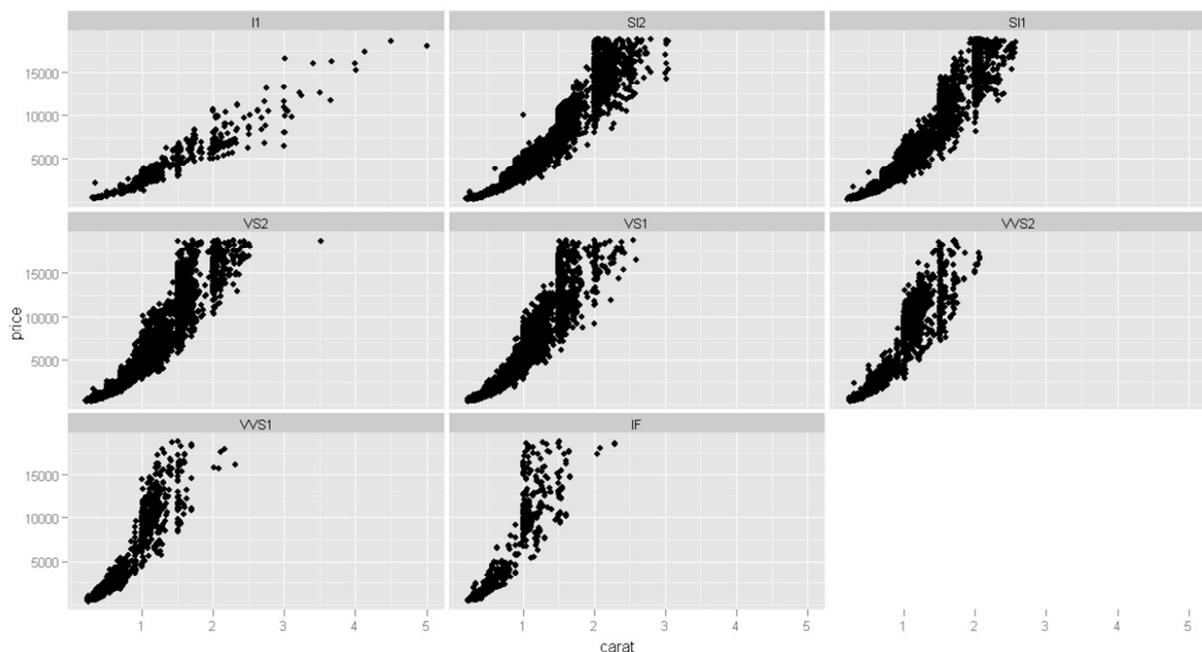


Abbildung 49: Beispiel einer *Small Multiples* Visualisierung<sup>38</sup>

Die Abbildung 49 zeigt eine *Small Multiples* Visualisierung des Preis-Karat Verhältnisses von Diamanten. Die Bilder zeigen die Korrelation des Preises von den Karat in Abhängigkeit der Reinheit des Diamanten. Die Diamanten steigen in Ihrer „Reinheit“ vom ersten Bild links oben „I1“ (bedeutet „imperfect 1“) bis zu intern „Lupenrein“ im letzten Bild unterste Reihe rechts „IF“ (bedeutet „internally flawless“) (Neubauer, 2012).

<sup>38</sup> Beispiel einer *Small Multiples* Visualisierung, anhand des Preis-Karat Verhältnisses von Diamanten. Auf der X-Achse ist der Preis aufgetragen und auf der Y-Achse der Karat-Wert. Somit bedeutet ein Anstieg von links unten nach rechts oben eine Wertsteigerung abhängig vom Karat-Wert des Diamanten.

Ebenfalls wurde die Klassifizierung von grafischer Darstellung, anhand der Datendichte einer Grafik vorgestellt. Die Datendichte ist die Anzahl von Einträgen in einer Datenmatrix, welche als Grafik angezeigt werden, die sich wiederum in der Datenmatrix befindet, in der die Beobachtungen als Variablen organisiert sind, getrennt durch die Fläche der Grafik. Es wurde beobachtet, dass die Grafiken sich in Ihrer Datendichte unterscheiden und dass Abbildungen die meisten Informationen beinhalten. Das Resultat von Tufte Analyse (Tufte, 1983 zitiert in Neubauer, 2012) war, dass Grafiken auf großen Datenmatrizen basieren sollten, da es immer besser ist zu viel Information anzubieten als zu wenig. Der Aufwand zusätzliche Informationen zu interpretieren und mit Ihnen umzugehen ist gering. Darstellungsdesigns mit wenigen Informationen, können zu dem Eindruck führen, dass Informationen fehlen oder wichtige Fakten ausgespart wurden. Daher sollte das Ziel sein Grafiken so zu designen, das die maximale Datendichte enthalten ist durch die Vergrößerung der Datenmatrix. Tufte präsentiert zwei Strategien um die Datendichte zu vergrößern. Die eine ist das Datenvolumen zu erhöhen wodurch die Form der Datenobjekte schrumpft, z.B. Verwendung von kleineren Punkten bei Scatter Plots. Die zweite Strategie ist die Reduzierung der Grafikfläche, dies ist das Konzept der *Small Multiples*, welches laut Tufte: ein effektives und gehaltvolles Grafikdesign ist (Neubauer, 2012).

Ein einfaches Beispiel für *Small Multiples* wäre es, von einem Spielfilm die einzelnen Bilder zu betrachten. Jedes Bild zeigt dieselbe Kombination von Variablen aber in unterschiedlichen zeitlichen Positionen. Dem Betrachter ist es möglich sich auf die Veränderung der Daten des Films über die Zeit zu konzentrieren. Ein weiteres Beispiel für die Umsetzung von *Small Multiples* wäre, die Entwicklung von verschiedenen Objekten, z.B. dargestellt als Datenpunkte, in kleinen Bildern. In einer derartigen Visualisierung kann die Veränderung der Objektdaten auf einen Blick verglichen werden. Auf einer online Seite der Firma Juce Inc. (Homepage der Firma Juce Inc., 2010 zitiert in Neubauer, 2012) listet deren CEO Zach Gemignani die folgenden drei Mindestvoraussetzungen für das designen von *Small Multiples* Visualisierungen. Der erste Punkt ist, dass die Anordnung der kleinen Bilder nach einer logischen Abfolge angelegt werden sollten, z.B. nach der Leseordnung von oben links nach unten rechts. Ansonsten wird der Betrachter von vielen kleinen und scheinbar nicht zusammenhängenden Bildern konfrontiert, welche schnell überfordern können und somit nicht für eine Analyse geeignet sind. Die zweite Voraussetzung ist es, die kleinen Darstellungen so klein wie möglich zu halten. Drittens sollten alle kleinen aufeinanderfolgenden aber nebeneinander angeordneten Bilder dieselbe Größe und Skalierung der Daten haben. Ein Beispiel hierfür ist die Abbildung 49, welche am Anfang dieses Exkurses zu finden ist (Neubauer, 2012).

Das Resultat von Tufte (Tufte, 1983 zitiert in Neubauer, 2012) war, dass *Small Multiples* wenn sie gut designt, verkleinert und mit hoher Grafikdichte umgesetzt wurden, einfach und effizient für Interpretation sind. Die Nachteile sind dabei die Limitierung der sichtba-

ren Anzeigengröße, zu viele Bilder überladen die Anzeige und verwirren, somit ist ein Überblicken der Gesamtentwicklung der Daten schwierig. Abschließend sind zusammengefasst die Vorteile von *Small Multiples*, dass UserInnen einen Überblick über die Daten erhalten und die Daten zu verschiedenen Zeitpunkten oder Daten von verschiedenen Objekten auf einen Blick vergleichen können. Diese Visualisierungstechnik auf jede andere Visualisierungstechnik angewandt werden, indem man einfach Bilder zu unterschiedlichen Zeitpunkten oder von unterschiedlichen Objekten verwendet.



