

# Entwicklung, Einsatz und Evaluation multimedialer Lernsoftware in der Statistikausbildung

Inauguraldissertation zur Erlangung des Grades  
eines Doktors der Wirtschaftswissenschaften  
der Fakultät für Wirtschaftswissenschaften  
der Universität Bielefeld

vorgelegt von

Dipl.-Kffr. Alexandra Pabst

Bielefeld, im Februar 2004

Dekan:

Prof. Dr. Rolf König

Gutachter:

Prof. Dr. Peter Naeve  
Prof. Dr. Joachim Frohn

## Inhaltsverzeichnis

<b>Einleitung - Motivation, Ziele, Vorgehensweise</b>	<b>1</b>
<b>I Statistik, Statistiker und Defizite in der Ausbildung</b>	<b>4</b>
<b>1 Statistik, Statistiker - eine erste Betrachtung</b>	<b>5</b>
1.1 Zitate, Meinungsäußerungen, Diskussionspunkte . . . . .	5
1.2 Statistik im täglichen Leben - Irrtümer . . . . .	10
1.3 Relevanz der Statistik für politische Entscheidungsprozesse . . . . .	16
<b>2 Statistikausbildung - Wandel und Defizite</b>	<b>18</b>
2.1 Kontroverse Meinungen . . . . .	18
2.2 Konstruktivismus, Ziele und neue Vermittlungsmethoden . . . . .	24
2.3 Statistical Bridges . . . . .	31
<b>3 Defizitklassifikation</b>	<b>33</b>
<b>4 Multimediale Lösungsvorschläge</b>	<b>37</b>
4.1 Methodenverständnis - illustriert an der Varianzzerlegung . . . . .	38
4.2 Simpson-Paradox – Kombination traditioneller Lehre und neuer Medien . .	51
4.3 Strukturen in Daten . . . . .	66
4.4 Problem der geeigneten Stichprobengröße . . . . .	77
<b>II Evaluation multimedialer Lernsoftware - das Fundament</b>	<b>88</b>
<b>5 Multimediale Lernsoftware - Definitionsversuche und Reflexionen</b>	<b>89</b>
<b>6 Evaluation - verschiedene Perspektiven</b>	<b>98</b>
6.1 Evaluation als Instrument - allgemeine Ansätze . . . . .	99
6.2 Evaluation von Multimedia - spezielle Ansätze . . . . .	103
<b>7 Konsequenzen für die Evaluationsstudie</b>	<b>111</b>
<b>III Prototyping - <i>Statistik interaktiv komplett</i></b>	<b>113</b>
<b>8 Phasenweise Entwicklung multimedialer Lernsoftware</b>	<b>113</b>
8.1 Prototyping als Methode der Softwareentwicklung . . . . .	113
8.1.1 Der evolutionäre Entwicklungsansatz . . . . .	113
8.1.2 Das gewählte Vorgehensmodell . . . . .	115
8.1.3 Projektziele und Prototypingziele . . . . .	120
8.1.4 Architekturentwurf . . . . .	120
8.2 Lernmodule von <i>Statistik interaktiv komplett</i> . . . . .	125
8.2.1 Multimediale Lernmodule und Komponenten . . . . .	126
8.2.2 Exkurs: Konzeption und Aufbau multimedialer Fallstudien . . . . .	128

8.3	Schwerpunkte und Ziele der Phasen des Prototypings . . . . .	133
8.3.1	Prototyp I - Architektur und Reportsplitter . . . . .	136
8.3.2	Prototyp II - statistische Perspektive . . . . .	137
8.3.3	Prototyp III - konträre Prinzipien . . . . .	141
8.4	Beschreibung einzelner Entwicklungsschritte des Statistiklabors . . . . .	144
<b>9</b>	<b>Erste Überprüfung der Lernmodule und des Einsatzkonzeptes</b>	<b>156</b>
<b>10</b>	<b>Kritische Reflexion zum Forschungsprojekt <i>Neue Statistik</i></b>	<b>159</b>
<b>IV</b>	<b>Explorative Evaluationsstudie - eine empirische Untersuchung</b>	<b>165</b>
<b>11</b>	<b>Einsatzszenarien, -konzepte und Evaluationsformen</b>	<b>165</b>
11.1	Charakterisierung: Einsatzkonzepte und Einsatzszenarien . . . . .	166
11.2	Assessment in Statistical Education . . . . .	168
11.3	Qualitative und quantitative Evaluation . . . . .	172
11.4	Varianten Ia und Ib des Einsatzkonzeptes . . . . .	172
<b>12</b>	<b>Explorative Evaluationsstudie - 1. Durchführung</b>	<b>175</b>
12.1	Rahmenbedingungen und Startschwierigkeiten . . . . .	175
12.2	Ausgewählte Komponenten des Einsatzkonzeptes Ia . . . . .	176
12.2.1	Statistische Motivations- und Theorievorlesung . . . . .	176
12.2.2	Statistische Reports . . . . .	178
12.2.3	Evaluationsinstrument Tutor - Tutorenmanagement . . . . .	180
12.2.4	Neue Prüfungsformen . . . . .	182
12.3	Fazit und Grenzen der qualitativen Evaluation . . . . .	185
<b>13</b>	<b>Explorative Evaluationsstudie - 2. Durchführung</b>	<b>187</b>
13.1	Das Einsatzkonzept Ib . . . . .	187
13.1.1	Komponenten und Ziele . . . . .	187
13.1.2	Integration eines Lernmanagementsystems . . . . .	190
13.1.3	Zielkompetenzen und Instrumente zur Zielerreichung . . . . .	195
13.1.4	Evaluationsinstrument als Standard . . . . .	196
13.2	Entwicklung und Einsatz des Evaluationsinstrumentes . . . . .	199
13.2.1	Zielsetzungen und strategischer Einsatz . . . . .	199
13.2.2	Untersuchungsschwerpunkte, Arbeitshypothesen, Evaluationsbögen .	201
13.2.3	Der Evaluationsreport . . . . .	203
<b>14</b>	<b>Ergebnisse und bleibende Erkenntnisse</b>	<b>219</b>
14.1	Gestaltung der Materialien . . . . .	219
14.2	Entwicklung des Einsatzkonzeptes . . . . .	220
	<b>Fazit und Ausblick</b>	<b>222</b>



---

<b>Literatur</b>	<b>227</b>
<b>I Bücher</b>	<b>227</b>
<b>II Zeitschriften</b>	<b>229</b>
<b>III Internetquellen</b>	<b>230</b>
<b>Anhang</b>	<b>231</b>
<b>A CD-Rom: Statistiklabor, Laborseiten, Bibliotheken</b>	<b>231</b>
<b>B Benutzeranweisungen zum Umgang mit Statistiklaborseiten</b>	<b>231</b>
<b>C Websites für Updates und weitergehende Informationen</b>	<b>232</b>
<b>D 1. Evaluationsbogen - Semesterbeginn</b>	<b>233</b>
<b>E 2. Evaluationsbogen - Semesterende</b>	<b>238</b>
<b>F Qualitative Befragung der 1. Durchführung</b>	<b>242</b>
<b>G Qualitative Befragung der 2. Durchführung</b>	<b>248</b>
<b>H Multiplechoiceklausur</b>	<b>252</b>
<b>I Technische Ausführungen zu Lernmanagementsystemen</b>	<b>254</b>
<b>J Zielkompetenzen und Instrumente zur Zielerreichung</b>	<b>256</b>
<b>K Diskussion der Arbeitshypothesen des Evaluationsinstrumentes</b>	<b>261</b>

## Abbildungsverzeichnis

1	Funktionen der Programmoberfläche . . . . .	23
2	Einbinden von Bibliotheken im Statistiklabor . . . . .	24
3	Defizitklassifikation für die Statistikausbildung . . . . .	36
4	Struktur des multimedialen Lösungsvorschlags . . . . .	40
5	Varianzzerlegung am Beispiel im Statistiklabor umgesetzt . . . . .	41
6	Theoriekomponente 1 umgesetzt in <i>Statistik interaktiv komplett</i> . . . . .	43
7	Theoriekomponente 2 umgesetzt in <i>Statistik interaktiv komplett</i> . . . . .	44
8	Theoriekomponente 3 umgesetzt in <i>Statistik interaktiv komplett</i> . . . . .	45
9	Vergleich der Stabdiagramme . . . . .	46
10	Idee der Varianzzerlegung . . . . .	47
11	Nichtidentifikation der Bimodalität . . . . .	49
12	Simpson-Paradox: Interaktivität im Statistiklabor . . . . .	55
13	Variationsmöglichkeit im R-Kalkulator . . . . .	56
14	Wahrscheinlichkeitstheoretische Überlegungen zum Simpson-Paradox . . . . .	58
15	Simpson-Paradox in einer 3x2-Kontingenztafel . . . . .	63
16	Variationsmöglichkeit im R-Kalkulator . . . . .	64
17	Histogramm des Gehirngewichtes . . . . .	66
18	QQNorm-Plot, Dichtefunktion und Boxplot . . . . .	69
19	Interaktivität: Durchführung verschiedener Transformationen . . . . .	70
20	Interaktivität: Wurzelziehung 2. Grades . . . . .	71
21	Gegenüberstellung zweier Transformationen . . . . .	71
22	Dichtefunktionen im Vergleich . . . . .	72
23	Interaktivität: Standardisierung der transformierten Daten . . . . .	73
24	Vergleich zweier Transformationen mit anschließender Standardisierung . . . . .	73
25	Interaktivität: Der gesamte Transformationsprozess . . . . .	74
26	Transformation: Ziehung 2ter Wurzel mit anschließender Standardisierung . . . . .	75
27	Transformation: Logarithmierung mit anschließender Standardisierung . . . . .	75
28	Dichtefunktionen der vier erhobenen Merkmale . . . . .	78
29	Dichtefunktion mit veränderter Fensterbreite . . . . .	79
30	Dichtefunktionen der drei Gruppen . . . . .	80
31	Demonstration des Gesetzes der großen Zahlen . . . . .	81
32	Lernmodul Erhebungsverfahren . . . . .	82
33	Verschiedene Stichprobenziehungen aus Fishers Irisdaten . . . . .	83
34	Variation der Stichprobenumfänge mit angezeigten Mittelwerten . . . . .	84
35	Gesetz der großen Zahlen - Gruppenmittel und Gesamtmittel . . . . .	85
36	Interaktivität: Gruppenmittel und Gesamtmittel . . . . .	86
37	Vorgehensmodell zum Prototyping von Sommerville . . . . .	115
38	Vorgehensmodell zum Prototyping . . . . .	116
39	Evolutionary Prototyping nach Sommerville . . . . .	118
40	Strukturmodell Ebene I . . . . .	121
41	Beziehungsmodell Ebene II . . . . .	122
42	Beziehungsmodell Ebene III . . . . .	123
43	Erweitertes Vorgehensmodell . . . . .	125
44	Exemplarische Startseite eines Falls . . . . .	130
45	Exemplarische Startseite einer Szene . . . . .	131
46	Szenen-Navigator . . . . .	132

---

47	Laborverknüpfungen in einer Szenenfallstudie . . . . .	132
48	Phasen I, II, III des Prototypings . . . . .	135
49	Coachingkomponente . . . . .	137
50	Eine Theoriekomponente des zweiten Prototypen . . . . .	140
51	Statistiklaborseite als Basis eines Reports . . . . .	142
52	Reporterstellung mit Hilfe der Berichterstellungsfunktion . . . . .	143
53	Statistischer Arbeitsplatz des Labors . . . . .	145
54	Visualisierungsmöglichkeiten des Grafik-Wizards . . . . .	146
55	Der Statistiktaschenrechner . . . . .	147
56	Dokumentation einer Statistiklaborseite . . . . .	149
57	Implementation a des Exkurses <i>Datenanalyse.R</i> . . . . .	151
58	Implementation b des Exkurses <i>Datenanalyse.R</i> . . . . .	152
59	Steckbrief zu den Funktionen . . . . .	152
60	Transformationen im Statistiklabor . . . . .	153
61	Texteditor als Laborobjekt . . . . .	154
62	Startseiten der Lernmodule - verschiedene Perspektiven . . . . .	157
63	Prototyp II: Appletbeispiel 1 . . . . .	162
64	Prototyp II: Appletbeispiel 2 . . . . .	163
65	Prototyp II: Appletbeispiel 3 . . . . .	163
66	Informationsströme zwischen den Beteiligten der Veranstaltung . . . . .	192
67	Die Entwicklung des Evaluationsinstrumentes . . . . .	198
68	Vierstufige Vorgehensweise . . . . .	200

*Ich stehe Statistiken etwas skeptisch gegenüber. Denn laut Statistik haben ein Millionär und ein armer Kerl jeder eine halbe Million.*

Franklin Delano Roosevelt (1882-1945)

## Einleitung - Motivation, Ziele, Vorgehensweise

**Motivation:** In einer Rede auf dem Joint Statistical Meeting in New York im August 2002 in betont der Statistiker und Kongressabgeordneter Tom Sawyer die Relevanz der Statistik für politische Entscheidungen. Er illustriert an einigen Beispielen, zu welchen Konsequenzen fehlende statistische Kenntnisse führen können. Seine Ausführungen verdeutlichen, wie wichtig es ist, Studierende, die später in politische oder wirtschaftliche Entscheidungsprozesse eingebunden sein könnten, mit einem soliden Verständnis für statistische Methoden und Konzepte auszustatten. Sawyer betont, dass ihnen zusätzlich zu statistischem Methodenwissen auch die Fähigkeit der kritischen Reflexion vermittelt werden muss.

Sawyer sieht ein großes Problem darin, dass zwar für viele politische Entscheidungen Daten existieren, aber niemand diese angemessen analysieren, geschweige denn interpretieren kann.

”The sad truth is that policymakers do not understand or use data well. There is often enough data available, sometimes even too much. But, the data are so poorly used, they drive policy in ways that mask rather than illuminate, confuse rather than classify, the issues at hand.”<sup>1</sup>

”We use an awful lot of data in Congress, but we don’t use it well. We craft initiatives and justify remedies in every conceivable societal arena. But, by large, we aren’t very sophisticated in our analysis, understanding, and applications of statistical information.”<sup>2</sup>

Zur Lösung dieses nicht zufriedenstellenden Istzustandes präsentiert Sawyer Vorschläge, von denen einige später betrachtet werden.

Die Zitate sollen dem Leser in einem ersten Schritt verdeutlichen, wie wichtig und aktuell die Statistik in der Realität ist und welche Konsequenzen mangelnde statistische Fähigkeiten haben können.

Die Arbeit stellt folgende These auf: Absolventen der bisherigen Statistikausbildung wären keine besseren *policymakers* als die von Sawyer Kritisierten. Von dieser These ausgehend untermauern Saywers Beispiele das Vorhaben dieser Arbeit: die Veränderung und Verbesserung der Statistikausbildung. Verbesserung in dem Sinne, dass versucht wird, bestehenden Defiziten entgegenzuwirken, Veränderung hinsichtlich der neuen Schwerpunktsetzung - mehr Gewicht auf das Verständnis statistischer Konzepte. Dies soll mit der Entwicklung und dem gezielten Einsatz multimedialer Lernsoftware erreicht werden.

**Ziele und Vorgehensweise:** Die vorliegende Arbeit setzt sich aus vier Teilen zusammen, die im Folgenden überblickartig beschrieben werden.

---

<sup>1</sup>[Sawyer2002], S. 9.

<sup>2</sup>[Sawyer2002], S. 9.

**Teil I: Statistik, Statistiker und Defizite in der Ausbildung**

Zuerst wird die Statistik als schwer, wichtig und vielschichtig charakterisiert. Dem Leser werden ausgewählte Zitate zur Statistik vorgestellt, Schwierigkeiten und Diskussionspunkte anhand der Literatur kritisch übermittelt. Es werden historische und aktuelle Artikel herangezogen, damit sich der Leser ein Bild von der zeitlichen Entwicklung der Diskussion um die Statistik und die Statistikausbildung verschaffen kann. Es wird die Konsequenz abgeleitet, dass die Statistikausbildung wichtig ist. Zudem wird aufgezeigt, inwiefern sich die Statistikausbildung innerhalb der letzten Jahrzehnte verändert hat: von traditioneller über rechnergestützte bis hin zur multimedialen Statistikausbildung. Der Schwerpunkt liegt bei den Defiziten der Statistikausbildung. Diese werden in den dargestellten Artikeln und Meinungsäußerungen unterschiedlicher Autoren deutlich. Es wird zudem gezeigt, dass die konstruktivistische Lerntheorie eine entscheidende Rolle bei der Neugestaltung der Statistikausbildung spielt. Anschließend wird eine Klassifizierung existierender Defizite vorgenommen.

Auf dieser aufbauend wird gezeigt, wie die neuen Medien einzelnen Defiziten positiv entgegenwirken können. Es werden exemplarisch erste Umsetzungen demonstriert: z.B. wird die Varianzzerlegung multimedial illustriert. Dies findet unter Anwendung der multimedialen Lernsoftware *Statistik interaktiv komplett* statt. Diese wurde im Forschungsprojekt *Neue Statistik* des BMBF<sup>3</sup> entwickelt, in dessen Rahmen die vorliegende Arbeit entstanden ist. Die konkreten Umsetzungen bringen den Leser in einem ersten Schritt mit der Lernsoftware in Kontakt und zeigen, wie neue Medien eingesetzt werden können, um die Statistikausbildung zu verbessern. Jeder Lösungsvorschlag setzt konstruktivistische Prinzipien um, zeigt, wie diese durch die neuen Medien in die Statistikausbildung integriert werden können.

**Teil II: Evaluation multimedialer Lernsoftware - das Fundament**

Der Teil II baut direkt auf der aufgestellten Defizitklassifikation auf, erklärt zunächst, was unter neuen Medien zu verstehen ist, was unter den vielseitig verwendeten Begrifflichkeiten *Multimedia* und *neue Medien* verstanden wird. Es wird außerdem der Versuch unternommen, den ebenfalls vielseitig verwendeten Begriff *Evaluation* zu definieren und herauszuarbeiten, dass Evaluation im Zusammenhang mit den neuen Medien eine entscheidende Rolle spielt. Die facettenreiche Diskussion um die Evaluation von Multimedia wird näher beleuchtet. Durch eine kritische Auseinandersetzung mit verschiedenen Ansätzen zur Evaluation wird die eigene Position verdeutlicht. Diese bildet das Fundament für das im Rahmen der Dissertation durchgeführte Evaluationsprojekt *Explorative Evaluation*, welches im Teil IV der vorliegenden Arbeit vorgestellt wird.

---

<sup>3</sup>Bundesministerium für Bildung und Forschung

**Teil III: Prototyping - *Statistik interaktiv komplett***

Eine Konsequenz der Betrachtungen in Teil II ist, dass der Entwicklungsprozess multimedialer Lernsoftware<sup>4</sup> bei der Bekämpfung der Defizite in der Statistikausbildung eine entscheidende Rolle spielt. Welche Aspekte bei dieser Entwicklung und Gestaltung zu beachten sind, wird in Teil III untersucht. Es wird gezeigt, dass die parallele Entwicklung konstruktivistischer Einsatzkonzepte für multimediale Lernsoftware notwendig ist, da der Einsatz Auswirkungen auf die Gestaltung der Lernsoftware hat. Ein Vorgehensmodell zum Prototyping multimedialer Lernsoftware wird aufgestellt und umgesetzt. Als Resultat entstehen die dem Vorgehensmodell gemäß schrittweise entwickelte und verbesserte Lernsoftware *Statistik interaktiv komplett* und ein konstruktivistisch geprägtes Einsatzkonzept für diese.

Die Ausführungen zeigen, wie konstruktivistische Ideen in die Gestaltung multimedialer Lernsoftware für die Statistikausbildung einfließen.

**Teil IV: Explorative Evaluationsstudie - eine empirische Untersuchung**

Im abschließenden Teil wird reflektierend über die durchgeführte Evaluationsstudie berichtet. Diese setzt sich aus zwei Durchführungen zusammen. Die Komponenten der Lernsoftware *Statistik interaktiv komplett* und das für diese entwickelte Einsatzkonzept werden in unterschiedlichen Versionen im Einsatz überprüft. Die schrittweise Verbesserung erfolgt dem in Teil III entwickelten Vorgehensmodell gemäß. In der ersten Durchführung der Evaluationsstudie wird qualitativ evaluiert, während in der zweiten quantitative und qualitative Evaluationsverfahren kombiniert werden. Zudem wird ein Evaluationsinstrument entwickelt und getestet, mit dessen Hilfe multimediale Lernsoftware sowie ihr Einsatz flächendeckend evaluiert werden können. Das Evaluationsinstrument bietet den Einsatz statistischer Analyseverfahren an, konzentriert sich somit auf die quantitative Evaluation. Die Umsetzung der Analyseverfahren findet auf speziell entwickelten Statistiklaborseiten statt. Das Statistiklabor ist eine zentrale Komponente der multimedialen Lernsoftware *Statistik interaktiv komplett*. Abschließend werden qualitative Elemente in das Evaluationsinstrument integriert, welche die quantitativen Evaluationsmethoden ergänzen. Die Notwendigkeit der Kombination qualitativer und quantitativer Evaluationsmethoden entsteht durch die speziellen Vor- und Nachteile beider einzelnen. Es wird außerdem gezeigt, unter welchen Umständen die Anwendung von Lernmanagementsystemen vorteilhaft für den Einsatz neuer Medien in der Statistikausbildung ist.

Die Argumentationsstränge der vier Teile bauen direkt aufeinander auf. In jedem Teil wird auf die in Teil I dargestellte Defizitklassifikation zurückgegriffen. Jeder Teil beginnt mit einer eigenen Einleitung, die die Ziele und die Argumentation des Teils darlegt. Dies soll die Übersichtlichkeit der Arbeit fördern und dem Leser Orientierung ermöglichen.

Es sei noch angemerkt, dass an den Stellen dieser Arbeit, an denen von Benutzern, Lernenden oder Lernern, Lehrenden oder Dozenten oder anderen Personengruppen die Rede ist, beide Geschlechter eingeschlossen werden. Um die Lesbarkeit der Texte nicht unnötig zu beeinträchtigen, wird von Differenzierungen abgesehen.

---

<sup>4</sup>Die Begriffe „Lernsoftware“ und „Lernsystem“ werden im Folgenden synonym verwendet.

## Teil I

# Statistik, Statistiker und Defizite in der Ausbildung

## Überblick über Ziele und Argumentation

Die Hauptaufgabe des Teils I besteht darin, eine Begründung für die Auswahl des Themas der Dissertation zu finden. Dies erfolgt durch den direkten Einstieg über das Problem: Es existieren Defizite in der Statistikausbildung. Den Ausgangspunkt bilden somit die Statistik und die Statistikausbildung. Um ein Fundament zu schaffen, von dem aus der Einsatz neuer Medien in der Statistikausbildung diskutiert werden kann, werden zuerst Fragen folgender Art aufgeworfen:

- Was ist Statistik?
- Warum ist Statistik wichtig? Warum ist die Statistikausbildung wichtig?
- Hat ein Wandel in der Statistikausbildung stattgefunden?
- Worin bestehen die Defizite der Statistikausbildung? Wie wurde bereits versucht, diesen entgegenzuwirken?
- Welche Defizite bestehen weiterhin? Sind diese zu klassifizieren?

Es wird schrittweise versucht, Antworten auf Fragen dieser Art zu geben und Sichtweisen verschiedener Statistiker darzulegen. Eine Auseinandersetzung mit kontroversen Positionen bezüglich der Statistik und der Statistikausbildung findet statt.

Teil I der Dissertation setzt sich aus den folgenden aufeinander aufbauenden Argumentationssträngen zusammen:

1. **Statistik, Statistiker - eine erste Betrachtung:** Die Erkenntnis lautet: „Statistik ist schwer und wichtig“. Wer früher nicht lesen und schreiben konnte, wurde betrogen. Heute wird der betrogen, der die in den Medien zahlreich vertretenen Grafiken und Statistiken nicht deuten und kritisch hinterfragen kann. Ausgewählte Zitate und Meinungsäußerungen vermitteln dem Leser zunächst ein Bild von der Statistik. Anschließend wird er mit Szenen aus dem Leben konfrontiert, in denen Statistikwissen hilfreich ist. Auf diesen aufbauend wird ein Beispiel diskutiert, das im Kontext der Politik anzusiedeln ist. Die Relevanz der Statistik für politische Entscheidungsprozesse wird herausgearbeitet. Der Leser wird in die Lage versetzt, auf Basis dieser Ausführungen eine erste Einschätzung und Charakterisierung der Statistik vorzunehmen.
2. **Statistikausbildung - Wandel und Defizite:** Aus der Erkenntnis „Statistik ist schwer und wichtig“ wird abgeleitet, dass die Statistikausbildung wichtig ist. Ihre historische Entwicklung und ihr Wandel werden dargelegt. Die Kombination von Statistik und Rechner im Sinne von *Computational Statistics* wird vorgestellt. Des Weiteren werden die Interdisziplinarität der Statistik sowie die Eigenschaften eines Statistikers diskutiert. Probleme der Statistikausbildung werden exemplarisch erläutert. Es wird Fragen der folgenden Art nachgegangen: Was sind die Grundpfeiler



der Statistikausbildung? Was sind die Prinzipien des Lernens von Statistik? Ziele der Statistikausbildung werden formuliert, neue Methoden vorgestellt und bewertet. Abschließend wird eines der Grundprobleme der Disziplin diskutiert - das Imageproblem.

3. **Defizitklassifikation:** Dieser Abschnitt ist eine Konsequenz, aber auch Zusammenfassung der voranstehenden Argumentationsstränge. Die Schwierigkeiten der Statistikausbildung werden zusammengestellt und klassifiziert. Die Defizitklassifikation bildet das Fundament für die folgenden Teile der Dissertation, die alle auf diese und die Bekämpfung der einzelnen Defizite abzielen.
4. **Multimediale Lösungsvorschläge:** Es werden Lösungsvorschläge entwickelt, einzelnen Defiziten durch den Einsatz der neuen Medien entgegenzuwirken. Mit Hilfe der statistischen Lernsoftware *Statistik interaktiv komplett*, insbesondere des Statistiklabors, werden statistische Methoden neu dargestellt, neu vermittelt. Die Vorteile des Statistiklabors werden an Beispielen demonstriert. Bei jedem demonstrierten Beispiel wird diskutiert, inwiefern welchen Defiziten entgegengewirkt wird.

## 1 Statistik, Statistiker - eine erste Betrachtung

Überblick: In diesem Kapitel wird der Versuch unternommen, die Statistik zu charakterisieren. Schwierigkeiten der Statistik und der Statistikausbildung werden anhand von Literatur diskutiert. Es entsteht ein Fundament für die abschließend zu entwickelnde Defizitklassifikation.

### 1.1 Zitate, Meinungsäußerungen, Diskussionspunkte

Es existieren unzählige Versuche, Statistik zu definieren. Diese unterscheiden sich teils stark, teils nur in Nuancen voneinander. Die Vielzahl der Versuche lässt erahnen, wie schwierig es ist, das Aufgabengebiet der Statistik klar zu beschreiben. Der folgende Blick auf Zitate dient dazu, einen ersten Eindruck zu gewinnen:

"Statistical thinking in the broad sense formed the silent background of the greatest scientific advantages of the 19th century."

"Statistical Science is the peculiar aspect of human progress which gives the twentieth century its special character." (Fisher, 1953) <sup>5</sup>

"... Statistics in its broadest sense is the matrix of all experimental science - and is consequently a branch of scientific method, if not Scientific Method itself; (...) The scientist should know Statistics as he knows logic and formal language for communicating his ideas." (Kendall, 1968) <sup>6</sup>

"Statistics is the science of gaining information from data! Data are numbers with contents." (Moore, 1995) <sup>7</sup>

Vardeman/Morris äußern sich kritisch über die Aufgabe der Statistik:

---

<sup>5</sup>[AndersonLoynes1987].

<sup>6</sup>[AndersonLoynes1987].

<sup>7</sup>[Moore1995].

"You know that statistics is about collecting and handling data. That is true, but incomplete; there is much more than that at work here. The vital point is that this discipline provides tools, patterns of thought, and habits of heart that will allow you to deal with data with integrity." (Vardeman/Morris 2003)<sup>8</sup>

"Honest statistical work has nothing to hide. It says what it says." (Vardeman/Morris 2003)<sup>9</sup>

Für Miron Straf<sup>10</sup> gehört zu der Frage *Was ist Statistik?* unmittelbar auch die Frage *Was ist ein Statistiker?*.

"Statisticians are those who have special expertise in any aspects of this practice, from developing theories to data mining, whatever may be their professional or disciplinary field." (Miron Straf 2002)<sup>11</sup>

Straf berichtet über eine Gruppe von Statistikern, die eine Definition für ihre Disziplin zu formulieren versuchen:

"As we stared transfixed at the words projected on the screen, someone from the back of the room meekly asked, 'Shouldn't the definition say something about data?'"<sup>12</sup>

Straf verdeutlicht anhand dieser Erfahrung, dass es nicht notwendig ist, sich auf eine Definition von Statistik zu einigen, um sehr gute statistische Arbeit zu leisten. Die verschiedenen Standpunkte, von denen aus Statistik betrachtet wird, stehen einander dabei nicht im Wege. Es wird deutlich werden, dass es für die Neugestaltung der Statistikausbildung wichtig ist, zentrale Aufgaben der Statistik festzulegen. Die zu entwickelnden multimediale Lernmodule müssen entsprechend gestaltet werden.

Betrachtet man die vielen Definitionsversuche zur Statistik, so könnte man zu dem Schluss kommen, dass Statistiker sich überhaupt nicht einig sind. Straf äußert dazu:

"But we do generally agree on the elements of which statistics is comprised: data; variability; uncertainty; sources of error; conceptualizing and quantifying phenomena; empirical inquiry through experiments, surveys, observational, and other studies; extraction and summary of information; analyses and inferences; and communication of results."<sup>13</sup>

Trotz dieser Übereinstimmung bezüglich genereller Aspekte gibt Straf zu bedenken, dass die Statistiker nie das volle Potential ihrer Disziplin erreichen oder ausschöpfen können, wenn sie nicht ein "common understanding of what statistics is about" teilen.

"How, for example, can we explain to public officials, the value of investments in statistics, when our field may appear to move in so many different, unrelated directions?"<sup>14</sup>

---

<sup>8</sup>[VardemanMorris2003], S. 21.

<sup>9</sup>[VardemanMorris2003], S. 26.

<sup>10</sup>ASA President 2002.

<sup>11</sup>[Straf2002], S. 2.

<sup>12</sup>[Straf2002], S. 2.

<sup>13</sup>[Straf2002], S. 2.

<sup>14</sup>[Straf2002], S. 2.

Der Leser wird bemerken, dass der interdisziplinäre Charakter der Statistik Schwierigkeiten hervorruft. Straf zitiert Michael Healy, der die Meinung vertritt, dass man die Statistik nicht als Wissenschaft, sondern eher als Technologie ansehen sollte, die einen fundamentalen und unersetzbaren Teil anderer Disziplinen darstellt. Er illustriert seine Position durch Forschungsbeispiele aus der Physik, Biologie, Wirtschaft und dem Ingenieurwesen, in denen die Statistik überall eine tragende Rolle spielt. Straf schließt mit den Worten:

”Statistics is important in industry, too, especially in the development of technologies; in public affairs; and in the education of our children, especially so that they understand variability and uncertainty, how to make sense from data, and the elements involved in making decisions.”<sup>15</sup>

Straf hat die Vision, der Statistik eine führende Position im wissenschaftlichen und technologischen Fortschritt einzuräumen. Er ist davon überzeugt, dass es gelingen wird, mit Hilfe der Statistik Informationen aus Daten zu gewinnen. Auf diese Weise kann die Statistik im öffentlichen Diskurs, der Politik wie auch bei der Ausbildung positiv mitwirken, indem sie vermittelt, wie mit Datenmaterial umgegangen wird, wie Entscheidungen unter Unsicherheit getroffen werden.

”It all starts with a common vision of what statistics is and what, together, we can achieve.”<sup>16</sup>

In der vorliegenden Arbeit wird die Lösung statistischer Probleme durch den Einsatz des Rechners im Vordergrund stehen. Die sich anschließende Dokumentation der Vorgehensweise und Ergebnisse in Form statistischer Reports wird als zentraler Teil der Statistik angesehen. Diese Aspekte bilden deshalb die Schwerpunkte der Entwicklung multimedialer Lernmodule.<sup>17</sup>

Die folgenden Ausschnitte aus dem Leitartikel der AMSTAT NEWS<sup>18</sup>, erschienen im Juli 2002 zeigen, dass die Diskussion um die Statistikausbildung bis zum jetzigen Zeitpunkt anhält und neue Wege gefunden werden müssen, Studierende für die Statistik zu gewinnen:

”The future of statistics lies in our future generations. We need to attract to our profession creative and energetic young people with quantitative talents. Yet, more and more of our high school youth are turned off by mathematics, and many of those entering college lack adequate quantitative skills. Before we blame high school mathematics instruction, we, as statisticians, must realize our own contribution to the problem.”<sup>19</sup>

Das Problem ist beschrieben. Wie diesem entgegengewirkt werden kann, wird im Folgenden genauer untersucht, einige Anregungen werden uns mit auf den Weg gegeben:

”What can we as statisticians, do? First, we need to attract students with quantitative talents from any discipline to graduate level statistics. We must develop multiple pathways to statistics. (...) We need to replace the filter of mathematics by a funnel whereby anyone with quantitative talents can enter statistics from any of a number of disciplines.”<sup>20</sup>

---

<sup>15</sup>[Straf2002] S. 3.

<sup>16</sup>[Straf2002], S. 3.

<sup>17</sup>Vgl. 10.

<sup>18</sup>Vgl. [Straf2002].

<sup>19</sup>[Straf2002].

<sup>20</sup>[Straf2002].

"We must see that children understand variability and uncertainty, how to make sense from data, and the elements involved in making decisions, not just because the concepts are important to our profession, but because they are important to our children in their future world. And we must see that all children learn them, because, with high quality education, all children are capable of learning them. In these basic statistical concepts, we must see that no child is left behind."<sup>21</sup>

Die Statistikausbildung legt den Grundstein dafür, wie Statistik durchgeführt oder angewendet wird, um Informationen zu gewinnen bzw. Probleme zu lösen, denn:

"We sometimes talk about statistics as though they were facts that simply exist, like rocks. (...) This is wrong. All statistics are created through people's actions: people have to decide what to count and how to count it, people have to do the counting and the other calculations, and people have to interpret the resulting statistics. (...) All statistics are social products, the results of people's efforts."<sup>22</sup>

Dieses Zitat benutzt den Begriff Statistik in doppelter Bedeutung: Statistik als Disziplin einerseits, als durch Menschen aufbereitete Daten andererseits.

Die ausgewählten Zitate und Diskussionen ermöglichen dem Leser einen kleinen Einblick in die vielseitige Diskussion um Statistik. Die Zitate verdeutlichen, wie schwierig greifbar die Disziplin und ihre Aufgaben sind.

Das folgende Unterkapitel vertieft diese ersten Eindrücke. Es wird gezeigt, in welchen Situationen statistische Kenntnisse eine Rolle spielen können. Zuvor allerdings wird eine etwas andere Art der Betrachtung von Statistik vorgestellt:

**"Pstatistics and Mstatistics"** Der zugrundeliegende Artikel<sup>23</sup> stammt von Ray Thomas<sup>24</sup> und ist im Mai 2003 als Leitartikel der RSS News erschienen. Thomas definiert die Statistik als eine notwendige Kombination aus "Pstatistics" und "Mstatistics" - eine Auffassung, die im Rahmen dieser Arbeit eine wichtige Rolle spielen wird.

Die beiden folgenden Zitate zeigen, welche Unterscheidung er vornimmt:

**Pstatistics:** " 'Statistics' in its plural meaning typically refers to facts about society. The charter of the Society refers to 'facts illustrating the condition and prospects of society in its material, social and moral relations, these facts being for the most part arranged in tabular forms (...) '. I'll call this plural meaning 'Pstatistics'."<sup>25</sup>

---

<sup>21</sup>[Straf2002].

<sup>22</sup>[Chan2002], S. 156.

<sup>23</sup>Vgl. [Thomas2003].

<sup>24</sup>Lifetime user of Pstatistics, honorary fellow of the Faculty of Social Sciences at the Open University, and the secretary of the Society's Official Statistics section committee.

<sup>25</sup>[Thomas2003], S. 1.

**Mstatistics:** "The Society is revising its charter in part because the word 'statistics' has changed its meaning. Developments since the 19th century have meant that 'statistics', used as a singular noun, has come to mean statistical methodology. I'll call this meaning 'Mstatistics', M standing for method and mathematics."<sup>26</sup>

Thomas kritisiert, dass Statistik in der Öffentlichkeit größtenteils auf Mstatistics reduziert wird: z.B. in der aktuellen Diskussion um Wartelisten von Krankenhäusern. Von Statistik wird erst dann geredet, wenn statistische Methoden auf die Daten angewendet werden. Vorher jedoch nicht, wenn über Wartezeiten und möglichen Schwindel bei der Datenerhebung diskutiert wird. Das Problem sieht Thomas darin, dass sich die beiden Teile der Statistik immer weiter voneinander entfernen und die methodische Auffassung von Statistik dominiert.

In der vorliegenden Arbeit stehen die Vermittlung der Statistik und die Neugestaltung der Statistikausbildung im Vordergrund. Wir schließen uns Thomas an, der die Auffassung vertritt, Studierende durch Pstatistics an die Disziplin heranzuführen, anstatt Statistik einseitig als methodisch zu charakterisieren.

"And if they become skilled and confident users of Pstatistics, are they not likely to become appreciative of the value of Mstatistics?"<sup>27</sup>

Im Rahmen der vorliegenden Arbeit wird dieses Prinzip umgesetzt. Herangeführt werden die Studierenden an die Statistik durch das Lösen realer Probleme. Probleme müssen beschrieben, eingegrenzt und gelöst werden. Während des statistischen Lösungsprozesses sollen die Studierenden die Notwendigkeit statistischer Methoden (Mstatistics) erkennen. Die Ergebnisse der Methodenanwendung sollen sie in einem sogenannten „statistischen Report“<sup>28</sup> (Pstatistics) dokumentieren und interpretieren. Sie lernen beide Seiten der Statistik gleichzeitig kennen und sollen während des statistischen Problemlösens die Notwendigkeit der Koexistenz beider verstehen. Thomas warnt:

"The Pstatistics idea of social or organizational construction is complementary to Mstatistics. It is too easy for statisticians to operate on numbers and pay only formal attention to the qualities of the Pstatistics they are using."<sup>29</sup>

Die in dieser Arbeit entwickelte Lernsoftware sowie ihr Einsatzkonzept streben die Kombination von P- und Mstatistics an. Exemplarische Umsetzungen werden im abschließenden Teil dieses Kapitels demonstriert und diskutiert. Auch in der traditionellen Lehre war eine Kombination von P- und Mstatistics möglich. Es wird die These aufgestellt, dass die Kombination dieser beiden Seiten der Statistik in den Lernmodulen besser zu implementieren ist.

**Zusammenfassung und Ausblick:** Es wurde in einem ersten Schritt gezeigt, wie schwierig es ist, die Statistik zu definieren. Der Aufgabenbereich der Statistik wurde durch Zitate beschrieben. Die Notwendigkeit der Kombination von P- und Mstatistics wurde diskutiert. Es wurde verdeutlicht, dass sie bei der Entwicklung multimedialer Lernsoftware für die Statistikausbildung eine entscheidende Rolle spielen wird.

---

<sup>26</sup>[Thomas2003], S. 1.

<sup>27</sup>[Thomas2003], S. 2.

<sup>28</sup>Dieser stellt eine zentrale Komponente der entwickelten Lernsoftware dar und wird dem Leser im Teil III dieser Arbeit detailliert vorgestellt.

<sup>29</sup>[Thomas2003], S. 3.

Die gewonnenen Eindrücke des Lesers werden im folgenden Unterkapitel vertieft. Irrtümer des täglichen Lebens und ihre Verbindung zu statistischen Kenntnissen werden aufgedeckt. Die Kombination der beiden Seiten P- und Mstatistics wird in einzelnen Beispielen aufgegriffen und näher erläutert werden. Anhand der Beispiele wird dem Leser demonstriert, dass die stärkere Betonung von Pstatistics in engem Zusammenhang mit der Integration konstruktivistischer Prinzipien in die neugestaltete Statistikausbildung steht.

## 1.2 Statistik im täglichen Leben - Irrtümer

Es werden Szenen des täglichen Lebens beschrieben, in denen die Statistik in unterschiedlichen Formen in Erscheinung tritt. Die Erkenntnis lautet: Statistik lauert überall. Statistische Kenntnisse können Irrtümer und Fehltritte im täglichen Leben vermeiden helfen. In den folgenden Szenen werden exemplarisch verschiedene Irrtümer diskutiert.

- 1. Szene: Messprobleme
- 2. Szene: Statistische Zusammenhänge versus Kausalzusammenhänge
- 3. Szene: Begrifflichkeiten und Interpretationen
- 4. Szene: Das arithmetische Mittel
- 5. Szene: Argumentation auf Basis von Häufigkeiten

Beginnen wir mit der Diskussion der einzelnen Szenen:

**Messprobleme**<sup>30</sup> - ein Thema, das uns in der Bundesrepublik schon einige Jahre beschäftigt. „Im ersten Quartal 2003 sind x Millionen Menschen arbeitslos“ - auf den ersten Blick erscheinen solche Zahlen, regelmäßig von der Bundesanstalt für Arbeit veröffentlicht, exakt zu sein. Schlagzeilen in den Medien wie „Arbeitslosigkeit viel höher als angegeben“ lassen vermuten, dass die Bestimmung einer exakten Zahl Probleme mit sich bringt. Dies kann zum einen auf die uneinheitliche Verwendung des Begriffs *Arbeitslosigkeit* zurückgeführt werden - die Bundesanstalt für Arbeit bezeichnet nur solche Personen als arbeitslos, die beim Arbeitsamt als arbeitssuchend gemeldet sind, mehr als 18 Stunden in der Woche arbeiten wollen, nicht nur vorübergehend Arbeit suchen, zwischen 15 und 65 Jahre alt sind und dem Arbeitsmarkt sofort zur Verfügung stehen. Was bedeutet dies? Habe ich es aufgegeben, mit Hilfe des Arbeitsamtes nach einer Anstellung zu suchen, oder suche ich eine Teilzeitarbeit, werde ich in der „exakten“ Zahl nicht erfasst.

Vergleicht man Arbeitslosenzahlen verschiedener Länder miteinander und klagt über die hohe Arbeitslosigkeit in Deutschland, muss man aufpassen: In einigen Ländern kann man gar nicht arbeitslos sein, wenn man vorher nicht beschäftigt war. Der klassische Hochschulabsolvent, der keine Anstellung findet, kann per Definition gar nicht arbeitslos sein. Bleiben wir innerhalb der Bundesrepublik, gibt es weitere Probleme: Übergangsarbeitslose oder Erwerbslose. Erwerbslos ist jeder, der gegen Entgelt arbeiten will, aber keine Stelle findet - auch die 80-jährige wohlhabende Dame, die sich selbst verwirklichen will. Betrachten wir die Definition des Arbeitsamtes, wird sie nicht als arbeitslos gezählt. Soviel zu den absoluten Zahlen. Die Arbeitslosenquote betrachtet relative Häufigkeiten: tatsächliche Arbeitslose, bezogen auf potentielle. Aber wer ist eigentlich potentiell arbeitslos, d.h. wie bilde ich diese Quote? Oder wie drehe ich geschickt am Nenner, um die Arbeitslosenquote

<sup>30</sup>Angelehnt an [Krämer1991], S.110ff.

zu senken oder zu erhöhen? Wächst der Nenner, teile ich also durch eine größere Zahl, wird der Bruch kleiner, die Arbeitslosenquote sinkt, ohne dass auch nur eine Person weniger tatsächlich arbeitslos ist. Eine schlagartige Reduktion der Arbeitslosenquote sollte stutzig machen. Auch die Vergleichbarkeit der Arbeitslosenquoten unterschiedlicher Länder ist damit zweifelhaft. In der Bundesrepublik besteht der Nenner aus „abhängigen Erwerbspersonen“. Selbstständige zum Beispiel umfasst diese Definition nicht. Warum sollten diese aber nicht als potentielle Arbeitslose gezählt werden? Nimmt man die Arbeitslosenquoten einiger Länder und bringt diese auf den gleichen Nenner, steht Deutschland gar nicht so schlecht da, wie immer behauptet wird.

Dieses Beispiel zeigt, dass ein geschärfter statistischer Blick, die Kenntnis über Datenerhebungen, Grundgesamtheiten, die Bildung absoluter und relativer Häufigkeiten, die Anwendung statistischer Kenntnisse auf reale Probleme eine kritische Meinung ermöglicht. Was bedeutet dies für die Statistikausbildung? Sie muss statistische Methoden realitätsnah vermitteln, den kritischen Blick der Studierenden schulen und auf mögliche statistische Fallen hinweisen. Ein reines Methodenverständnis kann hier nur als Grundlage betrachtet werden. Es handelt sich weniger um ein Methoden-, als um ein Messproblem und ein Problem des richtigen Vergleichs. Erinnern wir uns an die oben diskutierte Differenzierung zwischen M- und Pstatistics: die Kombination beider kann zu richtigen Ergebnissen führen.

### Statistische Zusammenhänge versus Kausalzusammenhänge

Der Schluss von Korrelation auf Kausalität ist einer der am weitesten verbreiteten Trugschlüsse. Von einem Kausalzusammenhang auf Korrelation zu schließen, ist immer richtig. Das Schwierige ist, dass es prinzipiell nicht zulässig ist, von Korrelation auf Kausalität zu schließen. Es bedarf immer noch einer zusätzlichen Begründung. Beispiele findet man in jedem Lehrbuch. Betrachten wir die positive Korrelation der Variablen *Ausländeranteil* und *Kriminalität*. Statistik kann missbraucht werden, um politische Entscheidungen herbeizuführen und Vorurteile zu unterstützen, indem von Korrelation der beiden Variablen auf Kausalität geschlossen wird. Großstädte wie Frankfurt oder Berlin ziehen sowohl Ausländer als auch Kriminelle an. Die positive Korrelation der beiden Variablen, der positive statistische Zusammenhang darf aber nicht als Kausalzusammenhang *mehr Ausländer, mehr Kriminalität* interpretiert werden. Die Größe der Städte sowie das Lebensalter wirken als nicht vernachlässigbare Hintergrundvariablen, die die positive Korrelation zumindest mitverantworten. Es wäre somit ein Trugschluss, von einem höheren Ausländeranteil auf höhere Kriminalität zu schließen, ohne den Einfluss weiterer Variablen zu untersuchen. Betrachten wir in diesem Zusammenhang ein weiteres Beispiel<sup>31</sup> und beginnen mit einem Abschnitt eines Zeitungsartikels der USA Today:

”Prayer can lower blood pressure” - ”Attending religious services lowers blood pressure more than tuning into religious TV or radio, a new study says. People who attend a religious service once a week and prayed or studied the Bible once a day were 40 % less likely to have high blood pressure than those who don’t go to church every week and prayed and studied the Bible less.” (Davis 1998).<sup>32</sup>

In dem Artikel wird fälschlicherweise darauf geschlossen, dass Beten und das Besuchen religiöser Messen einen niedrigeren Blutdruck verursachen. Die Unterstellung eines solchen Kausalzusammenhangs gehört zu einem Fehler, der in den Medien oft anzutreffen ist und

---

<sup>31</sup>Vgl. [Utts2003].

<sup>32</sup>[Utts2003], S. 75.

auf mangelnde statistische Kenntnisse schließen lässt.

Der im Artikel falsch dargestellte Kausalzusammenhang kann ebenso umgekehrt gelten: z.B. könnte es sein, dass gesunde Leute eher religiöse Veranstaltungen besuchen. Auch ist es denkbar, dass Leute, die stärker sozial engagiert sind, weniger gestresst sind und ohnehin einen niedrigeren Blutdruck haben. Es können viele mögliche Variablen gefunden werden, die den beschriebenen Zusammenhang erklären könnten. Utts zieht dieses Beispiel im *American Statistician* heran und fordert:

”Statistical studies are prominently featured in most major newspapers on a daily or weekly basis, yet most citizens, and even many reporters, do not have the knowledge required to read them critically. (...)

It is the responsibility of those of us teaching introductory statistics to make sure that our students are not among them.”<sup>33</sup>

Im Rahmen der vorliegenden Arbeit wird das Ziel verfolgt, die Statistikausbildung unter anderem so zu verbessern, dass Studierende statistische Methoden anwenden können und darüber hinaus die Fähigkeit besitzen, Ergebnisse statistisch kritisch zu interpretieren. Im zweiten Teil dieses Kapitels werden konkrete Vorschläge demonstriert, statistische Konzepte durch die neuen Medien besser als bisher zu vermitteln. An dieser Stelle kann der Bogen zum voranstehenden Unterkapitel gezogen werden: die stärkere Rolle von PStatistics in der Lehre kann den genannten Irrtümern vorbeugen. Dies wird auch durch das folgende Zitat deutlich:

”What good is it to know how to carry out a t-test if a student can not read a newspaper article and determine that hypothesis testing has been misused?”<sup>34</sup>

### **Begrifflichkeiten und Interpretationen**<sup>35</sup>

Schlägt man eine beliebige Tageszeitung auf, so begegnet einem das Wort *Prozent* im Durchschnitt über 100-mal. Mit Prozentangaben wird häufig versucht, Glaubwürdigkeit zu illustrieren. Außerdem zeigt es, dass man sich „mathematisch“<sup>36</sup> ausdrücken kann oder genaue Nachforschungen angestellt hat. Prozentzahlen sollen den Leser von einer bestimmten Meinung überzeugen, z.B. 60 Prozent der Befragten stützen diese Meinung. Befrage ich 10 Personen und 6 stützen diese Meinung, sind das genau 60 Prozent. Frage ich 10.000 Personen und 6000 stützen die Meinung, so sind das auch 60 Prozent. Diesen 60 Prozent vertrauen wir intuitiv mehr, erkennen, dass der Stichprobenumfang Auswirkung auf die Zuverlässigkeit der Daten, die Aussagefähigkeit des Ergebnisses hat. Doch haben Prozentzahlen auch ihre Berechtigung. Sicher sagen Prozentzahlen mehr aus als absolute Häufigkeiten. 10.000 Personen stimmten für Partei x. Für eine Kommunalwahl kann dies ein sehr gutes Ergebnis sein, für eine Bundestagswahl eine große Niederlage. Die absolute Zahl enthält keine Informationen über den Stichprobenumfang, eine Prozentzahl wäre viel aussagekräftiger gewesen.

Es ist hilfreich zu wissen, dass sich Prozentzahlen aus dem Quotienten aus absoluter Häufigkeit und Stichprobenumfang zusammensetzen. Der Statistiker spricht von der Bildung der relativen Häufigkeit und der anschließenden Multiplikation mit 100, um auf die Prozentzahl zu kommen. Die Information, aus welchen Zahlen die Prozentzahl gebildet wurde,

---

<sup>33</sup>[Utt2003], S. 78.

<sup>34</sup>[Utt2003], S.78.

<sup>35</sup>Vgl. [Krämer1991], S. 42ff.

<sup>36</sup>Sonderlich viel Mathematik kann der Statistiker allerdings nicht an der Bildung von Prozentzahlen entdecken.



wird verschluckt. 20 Prozent können sich aus  $1/5$ ,  $7/35$ ,  $117/585$ , ... zusammensetzen. Die auf diese Weise dargestellten relativen Häufigkeiten sagen mehr aus als die durch Multiplikation mit 100 entstehenden Prozentzahlen.

Schmückt sich ein Unternehmen damit, den Anteil der Frauen in Führungspositionen um 50 Prozent erhöht zu haben, so kann dies bedeuten, dass statt vorher 2 jetzt 4 Frauen in Führungspositionen tätig sind. Kleine Zahlen würden sich hinter einer beeindruckenden Prozentzahl verstecken und falsche Tatsachen vortäuschen. Das Ziel der Statistikausbildung sollte darin bestehen, den Studierenden Fähigkeiten zu vermitteln, aufbereitetes Datenmaterial kritisch zu betrachten. Das Beispiel zeigt, dass statistische Grundkenntnisse über absolute und relative Häufigkeiten und die Größe des Stichprobenumfangs ausreichen, um Prozentzahlen kritisch hinterfragen zu können. Die Aufgabe der Statistikausbildung muss darin bestehen, über das bloße Berechnen dieser Häufigkeiten hinaus (Mstatistics) inhaltliche Interpretationen und Bewertungen zu trainieren, auf Vor- und Nachteile absoluter und relativer Häufigkeiten aufmerksam zu machen (PStatistics).

**Das arithmetische Mittel** Im Alltag spricht man häufig vom *Durchschnitt* oder *Mittel* und meint das arithmetische Mittel. Es wird oft angewendet, weil es leicht zu berechnen ist und große Datensätze auf eine einzige Zahl komprimiert. Die Werte werden summiert und durch ihre Anzahl geteilt. Möchte man wissen, wieviel Geld Studierenden im Mittel monatlich zur Verfügung steht und befragt 10 Studierende, so kann das Ergebnis wie folgt aussehen:

250, 300, 300, 300, 300, 320, 320, 350, 350, 1500

Der Mittelwert betrage  $(250+300+300+\dots+1000)/10 = 397,2$  Euro

Der reine Mittelwert ohne die Berücksichtigung der Streuung ist nicht aussagekräftig aufgrund seiner Ausreißerempfindlichkeit. In vielen Situationen wird auf Basis des Mittelwerts argumentiert. Die Wahl scheint in einigen Fällen absichtlich auf den Mittelwert zu fallen, um Ergebnisse zu manipulieren.

„Der Begriff des Durchschnitts oder Mittels (Mittelwertes) - wie z.B. der Durchschnittslohn - wird oft als Synonym für den des Typischen gebraucht. Gibt es jedoch in einem Land nur einige außerordentlich reiche und zahlreiche sehr arme Familien, also sehr wenige riesengroße und sehr viele ganz kleine Einkommen, so ist das arithmetische Mittel aller Einkommen bei weitem nicht typisch. Der Median liefert ein viel typischeres Bild.“<sup>37</sup>

Der Median ist wie das arithmetische Mittel eine Maßzahl, die einen Datensatz auf *eine* Zahl komprimiert. Er teilt den sortierten Datensatz in zwei gleich große Hälften und nimmt den Wert in der Mitte zwischen diesen an. Deshalb reagiert er auf Ausreißer robust, d.h. er ist nicht ausreißerempfindlich wie das arithmetische Mittel. In dem Beispiel läge er zwischen den Werten 300 und 320, also bei 310 Euro. Es wird deutlich, dass er die tatsächlichen Daten besser als das arithmetische Mittel widerspiegelt.

Eine weitere Schwierigkeit besteht darin, dass von großen Teilen der Gesellschaft, z.B. in den Medien, das arithmetische Mittel fälschlicherweise als eine Art Normalzustand interpretiert wird: "im Mittel" und "normalerweise" werden gleichgesetzt. Dies ist z.B. in Davis, Kalifornien geschehen, wo ein Unternehmen versuchte, die Geruchsprobleme seiner Abwasseranlage auf ungewöhnlich starke Niederschläge zu schieben:

<sup>37</sup>[Szekely1990], S. 124.

”Last year’s severe odor problems were due in part to the extreme weather conditions created in the Woodland area by El Nino (according to a company official). She said Woodland saw 170 to 180 percent of its normal rainfall. Excessive rain means the water in the holding ponds takes longer to exit for irrigation, giving it more time to develop an odor.”<sup>38</sup>

Das Problem dieser Begründung besteht darin, dass der jährliche Niederschlag (Regenfall) stark streut. Betrachtet man die 5-Zahlen-Zusammenfassung<sup>39</sup> für Regenfall in inches für die Jahre 1951-1997, so fällt auf, wie groß der Range ist, d.h. wie groß das Intervall ist, das durch die Daten aufgespannt wird:

6.1, 12.1, 16.7, 25.4, 37.4

Der Statistiker spricht auch von einer großen Spannweite. Die Spannweite ist die Differenz aus dem größten und dem kleinsten Wert eines Datensatzes. Sie ist ein Streuungsmaß, das den Gesamtausdehnungsbereich der beobachteten Werte eines kardinalskalierten Merkmals in Bezug auf die Merkmalsachse angibt.

Im Jahr 1998, für das die Argumentation der Firma gilt, lag der Niederschlag bei 29.7 inches. Wir sehen anhand der 5-Zahlen-Zusammenfassung, dass dieser Wert zwar hoch, aber nicht außergewöhnlich hoch im Vergleich zu den anderen Jahren liegt. Der 5-Zahlen-Zusammenfassung kann man entnehmen, dass mehr als 75 Prozent der Jahre niederschlagsärmer als das Jahr 1998 waren, da der Wert von 29.7 inches zwischen dem 75 Prozentpunkt und dem Maximum liegt. Auch der Vergleich mit dem Median, der bei 16.7 inches liegt, spricht für hohe Niederschläge im Jahr 1998. Diejenigen, die aber mit ”excessive rain” oder ”abnormal” argumentieren, ziehen den Vergleich zum Mittelwert und übersehen, dass ein Regenfall von 29.7 inches als „noch normal” in Bezug auf den Range einzuordnen ist. Gerade bei Berichten über Niederschläge tritt dieser Fehler häufig auf. Es wird nicht verstanden, dass eine natürliche Variabilität besteht und eine Argumentation auf Basis des arithmetischen Mittels falsch ist.

Auch an diesem Beispiel wird deutlich, dass die Kombination von PStatistics und Mstatistics in der Statistikausbildung notwendig ist. Die bloße Vermittlung der beiden Lageparameter (Mstatistics) kann zu Interpretationsfehlern der eben beschriebenen Art führen.

Eine stärkere Betonung von Pstatistics (hier: der statistische sowie inhaltliche Vergleich der beiden Lageparameter) gibt den Studierenden nach Meinung des Autors die Fähigkeit, Maßzahlen kritisch zu bewerten und richtig einzusetzen.

**Argumentation auf Basis von Häufigkeiten** Das folgende Zitat zeigt, dass statistisches Verständnis eine notwendige Fähigkeit zu sein scheint, um sich im Leben seine eigene Meinung richtig zu bilden:

„Im Wetterbericht eines US-Fernsehsenders hieß es einmal, dass für Samstag die Regenwahrscheinlichkeit 50 Prozent betrage und die für Sonntag ebenfalls 50 Prozent. Der Meteorologe erklärte den Zuschauern abschließend, die Regenwahrscheinlichkeit für das kommende Wochenende betrage daher 100 Prozent.”<sup>40</sup>

---

<sup>38</sup>[Utts2003], S. 78.

<sup>39</sup>Die 5-Zahlen-Zusammenfassung ist die Zusammenfassung eines Datensatzes mittels Minimum, unterem Quartil (25 Prozentpunkt), Median, oberem Quartil (75 Prozentpunkt) und Maximum.

<sup>40</sup>[Gigerenzer2003], S. 41.

Neben der Wettervorhersage wird bei vielen z.B. wirtschaftlichen oder politischen Entscheidungen statistisches Verständnis benötigt. Es besteht ansonsten die Gefahr von Fehlentscheidungen, herbeigeführt durch Entscheidungsbefugte mit mangelnden Statistikkenntnissen.

Wir schon in einem der voranstehenden Beispiele diskutiert, stellt die Interpretation von absoluten/relativen Häufigkeiten eine Schwierigkeit dar. Es ist wichtig, dass eine eindeutige Bezugsmenge angegeben wird. Darauf aufbauend entstehen häufig Probleme, wenn von der Empirie, also von relativen Häufigkeiten auf die Modellebene, d.h. z.B. auf Wahrscheinlichkeiten gewechselt wird. Eine Wahrscheinlichkeit muss als relative Häufigkeit eines Ereignisses angegeben werden, die wiederum auf Basis eines großen Stichprobenumfangs ermittelt werden muss.

Betrachten wir ein Beispiel, in dem auf Basis von Häufigkeiten argumentiert wurde. Die zugrundeliegende Wahrscheinlichkeit wurde allerdings nicht nach den eben beschriebenen statistischen Prinzipien gebildet. Es wird eine Aussage über ein Risiko getroffen. Diese sollte statistisch korrekt durch empirische Häufigkeiten begründet sein, beruht aber fälschlicherweise auf subjektiven Meinungen:

Auf einem Plakat der DASA (Daimler Benz Aerospace) sind alle 94 bis dahin abgeschossenen Ariane-Raketen aufgeführt. Eine Ariane Rakete befördert Satelliten in ihre Umlaufbahn. Die Frage nach dem Risiko eines missglückten Starts wird im sogenannten Sicherheitsfaktor erfasst, dieser liegt bei etwa 99,6 Prozent. Das Plakat weist allerdings nur acht Fehlstarts aus.<sup>41</sup> Der statistisch geschulte Leser wird feststellen, dass ein Sicherheitsfaktor von 99,6 Prozent bei 94 gestarteten Raketen allerdings nicht durch acht Fehlstarts zustande kommen kann. Eine Nachfrage bei der DASA ergibt, dass der Sicherheitsfaktor nicht auf Basis der Fehlstarts berechnet wird, da diese menschliches Versagen beinhalten. Stattdessen werden die Konstruktionsmerkmale der einzelnen Teile der Rakete zur Berechnung herangezogen.

„(...), dass beispielsweise einer dieser Sterne letztlich auf ein Missverständnis zurückgehe, nämlich zwischen einem Arbeiter, der eine Schraube noch nicht montiert hatte, und seinem Kollegen von der nächsten Schicht, der annahm, sein Vorgänger habe die Schraube eingesetzt.“<sup>42</sup>

Zusammenfassend kann festgestellt werden, dass das Risiko eines Fehlstartes also nicht auf Basis beobachteter Häufigkeiten geschätzt wurde, sondern auf Grund der physikalischen Konstruktion der Maschinen. Diese Vorgehensweise ist nicht zwingend falsch, allerdings sollte das Zustandekommen der Prozentzahl in einem solchen Fall erläutert werden. Der Laie vermutet hinter einer Prozentzahl wie hier 99,6 Prozent eine aussagekräftige Größe, die auf Basis empirischer Häufigkeiten gebildet wurde, keine Größe, die tatsächlich aus subjektiven Meinungen resultiert. Es wäre hilfreich, den Zusammenhang zwischen der Realität und dem verwendeten Modell zu erklären: eine Konstruktionszeichnung kann als physikalisches Modell einer Maschine aufgefasst werden. Auf Basis dieses Modells kann das Risiko der einzelnen Teile der Maschine abgeschätzt werden. Es ergibt sich additiv das Gesamtrisiko eines Fehlstarts dieser Maschine, wobei Interaktionen von dem physikalischen Modell nicht unbeachtet bleiben dürfen.

<sup>41</sup>Pro Fehlstart wird ein Stern abgebildet.

<sup>42</sup>[Gigerenzer2003], S. 48.

**Fazit:** Es existieren zahlreiche Situationen im täglichen Leben, in denen statistische Paradoxe auftreten oder Fehlentscheidungen aufgrund mangelnder statistischer Kenntnisse zustande kommen. Die in diesem Abschnitt vorgestellten Beispiele aus dem Leben sollten dem Leser einen Einblick in die Problematik verschafft haben. Das Ziel der Statistikausbildung sollte darin bestehen, Studierende auszubilden, die Situationen der vorgestellten Art kritisch bewerten und auf Basis ihrer M- wie auch Pstatistics-Kenntnisse in der Lage sind, Fehlentscheidungen zu vermeiden.

### 1.3 Relevanz der Statistik für politische Entscheidungsprozesse

Dieses Unterkapitel hat das Ziel, dem Leser die Relevanz statistischer Kenntnisse in politischen Entscheidungsprozessen vor Augen zu führen. Es wird außerdem gezeigt, welche Defizite auf diesem Gebiet zur Zeit bestehen. Dazu wird der zu Beginn dieser Arbeit angeführte Artikel des Statistikers Sawyer wiederholt aufgegriffen. Sawyer beschreibt Schwierigkeiten, die in der Politik durch fehlende statistische Kenntnisse auftreten. Diese werden als Ansatzpunkte und Begründungen dafür verwendet, die Statistikausbildung zu verbessern. Viele von Sawyer angesprochenen Aspekte werden in die entwickelte Defizitklassifikation für die Statistikausbildung aufgenommen.<sup>43</sup>

Sawyer berichtet über die Schwierigkeiten des amerikanischen Kongresses, vorhandenes Datenmaterial effizient auszuwerten.

”It represents a way to bridge the gaps between science and policy. One of the greatest barriers to informed and prudent decision making is the appalling ignorance of the real meaning and uses of data.”<sup>44</sup>

Diesem nicht zufriedenstellenden Istzustand setzt Sawyer zahlreiche Lösungsvorschläge entgegen. Im Folgenden werden die Vorschläge näher betrachtet, die sowohl relevant für die Charakterisierung der Statistik als auch für die Herausarbeitung existenter Defizite in der Statistikausbildung sind; ”we must do a better job of using your science to”.<sup>45</sup>

**Determine what we can realistically measure, and the limitations of those measurements.** Sawyer gibt ein Beispiel aus der Bildungspolitik. Dort wurden Entscheidungen über milliardenschwere Bildungsprogramme auf Basis von 14 Jahre alten Daten getroffen. Dies führte zu einer Reihe bildungspolitischer Fehlentscheidungen, die dadurch gelöst wurden, dass Bildungspolitiker, Statistiker und Erziehungswissenschaftler gemeinsam einen Lösungsvorschlag erarbeiteten. Einmal im Jahrzehnt wurden Daten bezüglich des Einkommens/der Armut der Bevölkerung erhoben. 1980 erhobene Daten (diese geben das Bevölkerungseinkommen des Jahres 1979 wider) wurden noch im Jahr 1993 benutzt, um den Bedarf der einzelnen Gemeinden an staatlicher Förderung der Bildung zu bestimmen. Mit anderen Worten: Entscheidungen über milliardenschwere Bildungsprogramme wurden auf Basis veralteter Daten getroffen. Die falsche Verteilung der Ressourcen hatte Konsequenzen:

”Our efforts to target assistance to educationally disadvantaged children were further skewed by a lack of precise data.”<sup>46</sup>

---

<sup>43</sup> Vgl. 3.

<sup>44</sup> [Sawyer2002], S. 9.

<sup>45</sup> [Sawyer2002], S. 9.

<sup>46</sup> [Sawyer2002], S. 9.

Sowohl die Relevanz als auch der interdisziplinäre Charakter der Statistik werden hier deutlich.

”The solution wasn’t perfect from a policy perspective, or easy from a scientific one, but it demonstrated how statistics could play an enormously constructive role in achieving policy objectives.”<sup>47</sup>

Die Vermittlung der Fähigkeit, an solchen Entscheidungsprozessen mitwirken zu können, sollte eines der übergeordneten Ziele der Statistikausbildung sein. Dabei geht es nicht nur darum zu erkennen, dass veraltetes Datenmaterial zu Fehlentscheidungen führen kann. Ein weiteres Ziel kann darin bestehen, die Art der Datenerhebung kritisch zu hinterfragen. Dazu müssen die Existenz, Vor- und Nachteile der verschiedenen Methoden der Datenerhebung und der Stichprobenziehung vermittelt werden. Um z.B. das Durchschnittseinkommen der US-Bevölkerung zu ermitteln, ist ein einstufiges Verfahren nicht praktikabel. Stattdessen kann man effizienter vorgehen, indem man eine Klumpenstichprobe zieht: die interessierende Grundgesamtheit zerfällt in natürliche Klumpen, z.B. Bundesstaaten bzw. Gemeinden. Die Klumpenstichprobe bietet sich an. (Eine geschichtete Zufallsstichprobe würde je nach Schichtungsvariable unnatürliche Schichten bilden.) Auf der ersten Stufe dieser Art der Stichprobenziehung werden nur einige der Klumpen zufällig ausgewählt, die erste Stufe ist demnach eine Teilerhebung. Die zweite Stufe ist eine Vollerhebung, da alle Merkmalsträger der gewählten Klumpen in die endgültige Stichprobe gelangen. Damit handelt es sich um eine unechte, zweistufige Stichprobenziehung. Dieses Beispiel demonstriert, dass es bei der Statistikausbildung insbesondere um die zusätzliche Vermittlung kritischer Einschätzungen der Methoden gehen muss. In diesem Fall kann dies bedeuten, dass der Entscheidungsbefugte die Vor- und Nachteile verschiedener Datenerhebungen/Stichprobenziehungen kennen muss. Er muss z.B. wissen, worin der Vorteil der Klumpenstichprobe gegenüber der geschichteten Zufallsstichprobe besteht, was gemeint ist, wenn von dem *Schichtungseffekt*<sup>48</sup> oder *Klumpeneffekt*<sup>49</sup> die Rede ist. Man erkennt die schon mehrmals demonstrierte Notwendigkeit von P- und Mstatistics.

**Analyze, present and explain statistical results in ways that are meaningful to data users in the policy arena** - Sawyer vertritt die Meinung, Statistiker mehr in politische Entscheidungsprozesse einzubinden, um die richtigen Fragen zu stellen, somit die richtigen Daten zu erheben und dabei zu wissen, welche statistischen Methoden existieren und wie sie angewendet werden können, um Antworten zu finden. ”We need to do a better job of interpreting data. We need to cut through statistical noise so that legislative solutions that rely on data become laser-guided missiles - more likely to hit their target with precision. We need to help to do that.”<sup>50</sup> Auch wenn diese Aussage amerikanisch und militärisch formuliert ist, trifft sie den Punkt: es werden Statistiker gebraucht, um sowohl die Qualität der Daten als auch die der statistischen Analysen und Interpretationen zu sichern, denn auf deren Basis werden wichtige politische Entscheidungen getroffen. Sawyer betont, dass in Zukunft fundamentale statistische Kenntnisse gebraucht werden, um Daten und somit Informationen effizient zu managen und zu analysieren. Dafür werden

---

<sup>47</sup>[Sawyer2002], S. 9.

<sup>48</sup>In der Regel ist die Schätzung für die interessierende Größe genauer als im Fall einer einfachen Zufallsstichprobe, da in der geschichteten Zufallsstichprobe die Struktur der Grundgesamtheit besser wiedergegeben wird.

<sup>49</sup>Der sogenannte Klumpeneffekt beschreibt den Auswahlfehler, den man im Vergleich zur einstufigen Stichprobe macht, wenn die Klumpen schlecht gewählt sind.

<sup>50</sup>[Sawyer2002], S. 12.

Statistiker oder z.B. Wirtschaftswissenschaftler mit einer fundierten Statistikausbildung gebraucht.

**Build confidence in data collection, analysis, and publication, by preserving data confidentiality and protection respondent privacy** - Dieser Aspekt findet sich in der in dieser Arbeit aufgestellten Defizitklassifikation<sup>51</sup> unter *fehlendem Image der Statistik und ihrer allgemeinen Unterschätzung* wieder. Um diesem Defizit entgegenzuwirken, muss u.a. sichergestellt werden, dass die Bevölkerung Vertrauen in die Statistik hat. Sawyer argumentiert an einem aktuellen Beispiel:

”Overall, I believe, Congress acted appropriately after the terrorist attacks of September 11. We did the best we could, based on what we knew. But we had neither the time nor the knowledge to make fully informed decisions in all cases. And we made at least one serious error that we ought to undo.”<sup>52</sup>

Dieser besteht darin, dass statistische Daten, die vertrauliche, persönliche Informationen enthalten und ursprünglich nur für statistische Zwecke erhoben wurden, nun für die Fahndung möglicher Terroristen benutzt werden. Solche Aktionen führen dazu, dass das Misstrauen gegenüber der Statistik wächst, Fragebögen nicht mehr ausgefüllt werden, es also immer schwieriger wird, an Informationen zu kommen. Da diese die Basis vieler politischer Entscheidungen bilden, werden auch diese fehlerhaft. Um das Image der Statistik zu verbessern, muss in Zukunft die Integrität der Statistik gewahrt werden.

**Zusammenfassung:** Die Ausführungen haben verdeutlicht, dass die Statistik relevant für politische und wirtschaftliche Entscheidungsprozesse ist. Die dargestellten Konsequenzen mangelnder statistischer Fähigkeiten untermauern das Vorhaben dieser Arbeit, die Statistikausbildung zu verändern und zu verbessern. Bevor jedoch von einer Veränderung oder Verbesserung die Rede sein kann, werden Ziele, Mittel und Konsequenzen der bestehenden Statistikausbildung in Ansätzen beschrieben.

## 2 Statistikausbildung - Wandel und Defizite

### 2.1 Kontroverse Meinungen

Dieses Kapitel hat die Aufgabe, dem Leser die bestehende Statistikausbildung in Ansätzen zu beschreiben. Insbesondere werden die Defizite und der durch sie ausgelöste Wandel der Statistikausbildung dargestellt. Der Übergang von der traditionellen zur rechnergestützten Statistikausbildung wird kurz beschrieben, Reformvorschläge diskutiert. Anschließend werden Ziele der Statistikausbildung und Methoden zu ihrer Neugestaltung beschrieben und bewertet. Abschließend erfolgt die Diskussion eines 2002 erschienenen Artikels zur Neugestaltung der Statistikausbildung, der zeigt, dass die Statistik weiterhin ein Imageproblem hat.

---

<sup>51</sup>Vgl. 4.4.

<sup>52</sup>[Sawyer2002], S. 12.

### Statistikausbildung in der Diskussion

In der Grundausbildung der Wirtschafts- und Sozialwissenschaften<sup>53</sup> gilt das Fach Statistik seit jeher als eher unpopulär und/oder uninteressant, wird in der Regel nicht als wichtiger Teil der Ausbildung betrachtet und von einem Großteil der Studierenden als Hürde angesehen. Als wesentliche Gründe für die Probleme der Studierenden mit der Statistik werden die Nähe zur Mathematik, sowie der fehlende Bezug zum eigenen Studium und späteren Berufsleben genannt. Die obigen Ausführungen haben dem Leser gezeigt, dass die Statistik sowohl in der Wirtschaft, als auch in der Politik eine entscheidende Rolle spielt. Beide Bereiche sind typische Berufsfelder eines Wirtschaftswissenschaftlers.

Betrachtet man Studierende, die am Ende ihres Studium stehen, genauer, so trifft man häufig auf solche, die nur über ein oberflächliches Statistikwissen verfügen. Ihnen fehlt das tiefer gehende Verständnis bestimmter Sachverhalte, die Fähigkeit, Probleme als solche zu erkennen, Zusammenhänge zu sehen erlernte Kenntnisse gezielt anzuwenden.

Um diesem Zustand entgegenzuwirken wird seit einigen Jahren seitens deutscher Hochschulen an einer Neugestaltung der statistischen Grundausbildung gearbeitet. Die Intention dieser Neugestaltungsgedanken besteht prinzipiell darin, das Interesse der Studierenden für die Statistik zu wecken. Sie sollen darauf aufmerksam gemacht werden, dass ihnen fundierte Statistikkenntnisse in vielen Bereichen des Lebens behilflich sein können. Ein weiteres Ziel besteht darin, den Studierenden in einer neu gestalteten Statistikausbildung zu verdeutlichen, dass die Statistik keineswegs als Ableger der Mathematik zu verstehen ist, sondern in vielen großen Disziplinen eine entscheidende Rolle spielt. Der interdisziplinäre Charakter der Statistik wird stärker betont.

Fisher formulierte das Ziel der Statistikausbildung wie folgt:

„Der Studierende soll durch das Studium des Verfahrens befähigt werden festzustellen, auf welche Fragen, die sich aus seinem eigenen Material ergeben, ein solches Verfahren imstande ist, eine bestimmte Antwort zu geben, und festzustellen, was für weitere Beobachtungen nötig wären, um andere wichtige Fragen zu lösen.“<sup>54</sup>

Es fällt auf, dass Fischer vor Jahrzehnten nichts anderes forderte als Saywer heute: eine Kombination von P- und Mstatistics in der Statistikausbildung, um den Studierenden Problemlösungskompetenz zu vermitteln.

**Computational Statistics** Im Rahmen einer Neugestaltung der Statistikausbildung findet seit einigen Jahren eine Diskussion über den Einsatz und die Qualität neuer Entwicklungen auf dem Gebiet der Medien- und Informationstechnik statt. Der Aspekt der neuen Medien wird noch einen Augenblick lang ausgeklammert. Es werden zunächst Aspekte berücksichtigt, die im Zusammenhang mit der rechnergestützten Statistikausbildung diskutiert werden.

Zunächst fällt auf, dass sich mit der Benutzung des Computers der Rechenaufwand des Studierenden erheblich reduziert. Dadurch gewinnt er Zeit, sich mit anderen Aspekten als der bloßen schematischen Berechnung zu beschäftigen. Da diese Berechnungen nun zur Nebensache werden, kann sich der Student verstärkt auf die Interpretation und Analyse

---

<sup>53</sup>In dieser Arbeit rückt die Statistikausbildung der Wirtschaftswissenschaftler in das Zentrum der Betrachtungen.

<sup>54</sup>[Fisher1956].

der vom Computer gelieferten Ergebnisse konzentrieren.

Wichtiger als der verminderte Aufwand ist allerdings die Tatsache, dass der Einsatz von Computern in der Statistikausbildung eine sehr praxisbezogene, realitätsnahe Arbeit ermöglicht: mit dem Computer ist man in der Lage, große Datensätze zu analysieren und Realitäten zu simulieren. Es wird betont, dass Studierende mehr Interesse an theoretischen Methoden zeigen, wenn sie ihnen anhand von realistischen Situationen vermittelt werden. Diese umfassen selten eine so geringe Anzahl von Daten, dass sie per Hand zu analysieren wären. Hier kommt der Vorteil des Computers zum Ausdruck, der mit großen realen Datensätzen umgehen kann.

Die Erfahrungen der rechnergestützten Statistikausbildung zeigen, dass der Einsatz von Computern kritisch betrachtet werden muss. Studierende neigen dazu, zu berechnen, was zu berechnen ist, ohne die angewendeten Methoden kritisch zu hinterfragen, z.B. zu überprüfen, ob die erzielten Ergebnisse dazu dienen, dem Ziel ihrer Arbeit näher zu kommen. Diese Vorgehensweise wird durch die Benutzung eines Computers extrem verstärkt, da Methoden auf Knopfdruck angewendet werden können, die Anwendung schnell und einfach erscheint.

In der rechnergestützten Lehre hat der Lehrer die Aufgabe, den Studierenden den Rechner als Instrument zum Erlernen von Statistik näher zu bringen. Er sollte versuchen, sie mit dem Computer zum Lernen zu motivieren, indem er ihnen die Möglichkeiten des Rechners aufzeigt, z.B. die Leichtigkeit, Daten geringfügig zu variieren und das Resultat graphisch zu betrachten. Der Lehrer muss verhindern, dass die Studierenden den Rechner allein zur Automatisierung ihrer Rechenschritte benutzen.

David S. Moore äußerte sich wie folgt zu diesem Thema:

"I also feel strongly that heavy use of computing technology is essential for realistic learning of practical statistics (...) I recognize that computing can frustrate students. Real statisticians use SAS<sup>55</sup> is not a reason to use SAS in a first course!"<sup>56</sup>

Durch die Einbeziehung des Rechners in die Statistikausbildung verändern sich auch deren Inhalte. Methoden können anhand realer Datensätze erklärt werden, Simulationen in die Lehre einbezogen werden. Der Dozent hat die Möglichkeit, größeres Interesse durch Verwendung aktueller Datensätze und ihre gezieltere Analyse und Interpretation zu wecken. Scheaffer dazu:

"The increasing use of computers, not just within the discipline but in society in general, has placed an increasing premium on quantitative reasoning in general and on statistical reasoning in particular (...) It is already apparent that debates about public policy issues tend to involve reasoning with data."<sup>57</sup>

Es existieren unterschiedliche Schwerpunktsetzungen. Einige Vertreter setzen den Rechner ein, um das Hantieren mit großen Datensätzen zu demonstrieren und Datenanalysen zu veranschaulichen. Andere führen die Studierenden über den Rechner an die Programmierung heran, diskutieren erstellte Funktionen und lassen die Studierenden schließlich selbst erste Funktionen/Programme schreiben. Diese unterschiedlichen Möglichkeiten entstehen auch beim Einsatz der neuen Medien in der Statistikausbildung. Vor der Diskussion der

---

<sup>55</sup>Statistisches Analysesystem: Datenverarbeitung und statistische Auswertung

<sup>56</sup>[Moore1997].

<sup>57</sup>[Scheaffer2002], S. 6.



neuen Medien werden einige Ausführungen David S. Moores zur Neugestaltung der Statistikausbildung betrachtet.

### **New Pedagogy and New Content - David S. Moore**

Moore stellt die Frage nach einer neuen Form der pädagogischen Gestaltung der Statistikausbildung. Sein Hauptaugenmerk liegt auf der Umsetzung konstruktivistischer Prinzipien. Er plädiert für das Zusammenspiel von neuem Inhalt, neuer Pädagogik und neuer Technologie. Dabei repräsentiert der Inhalt den Lehrstoff, die Pädagogik die Maßnahmen, die ergriffen werden, um den Studierenden das Lernen zu erleichtern. Die Technologie steht für Computer, Kommunikation und Multimedia. Moore betont, dass die Technologie sowohl den Inhalt, als auch die Art des Unterrichts entscheidend beeinflusst. Die ausgewogene Mischung der drei genannten Bereiche ermöglicht effektives Lernen. Dieses sollte besonders darauf ausgerichtet sein, bedeutungsvolle Prinzipien zu verinnerlichen und tiefer gehendes statistisches Verständnis zu gewinnen.<sup>58</sup> Erforderlich für die Erreichung dieses Ziels ist das aktive Lernen der Studierenden: sie müssen sich ihr Wissen selbst konstruieren und organisieren, indem sie neu gewonnene Erfahrungen und Ideen mit bereits gefestigten Ansichten und Konzepten kombinieren. Dies ist eines der Grundprinzipien des Konstruktivismus. "Learning how to learn" lautet die neue Devise, nach der ausgebildet wird. Der Lehrer tritt nur am Rande als eine Art Moderator, führende Hand in Erscheinung. Moore stellt seine "reform of pedagogy" vor.<sup>59</sup>

- **Goals:** Higher-order thinking, problem solving, flexible skills applicable to unfamiliar settings.
- **The new model:** Students learn through their own activities; a good teacher encourages and guides their learning.<sup>60</sup>
- **What helps learning:** Group work in and out of class; explaining and communicating; frequent rapid feedback; work on problem formulation and open-ended problems.

Konstruktivistische Prinzipien erhalten Einzug in die Statistikausbildung. Die Tendenz geht weg vom Lehrer, der versucht, sein Wissen durch reines Vortragen zu vermitteln. Es entstehen dynamische Unterrichtsformen, in der der Schüler durch seine aktive Teilnahme den Prozess des Lernens und des Verstehens vorantreibt. Diese neue Form ist gekennzeichnet durch Diskussionen innerhalb der Gruppe, gemeinsames Lösen statistischer Probleme anhand aktueller Themen und Daten und Problemlösen mit Hilfe neuer Technologien. Doch nicht nur die Art des Unterrichts, sondern auch sein Inhalt verändert sich. Moore hält zum Beispiel formale Beweise, die viel Mathematik in sich bergen, oder auch die formale Seite der Wahrscheinlichkeitsrechnung für wenig hilfreich bezogen auf Anfängerkurse der Statistik. Viel wichtiger ist die Analyse kleiner Datensätze, an denen die Studierenden die Nähe zur Realität erkennen, das Erzeugen von Grafiken und insbesondere die ihre Interpretation.

---

<sup>58</sup>Wieder erkennen wir die Notwendigkeit der Kombination von P- und Mstatistics.

<sup>59</sup>Vgl. [Moore1997].

<sup>60</sup>In Abgrenzung zu "The old model: Students learn by absorbing information; a good teacher transfers information clearly and at the right rate."

"(...) the formal machinery is a barrier rather than a help at this level!"<sup>61</sup>

"Formal probability does not help these students master the ideas of inference, and it depletes reserves of mental energy that might better be applied to essentially statistical ideas."<sup>62</sup>

Aktives Lernen, guided learning, die neue Rolle des Lehrers - dies sind Elemente des Konstruktivismus, die schrittweise Einzug in die Statistikausbildung erhalten haben. Der Konstruktivismus spielt auch bei der Gestaltung der neuen Medien, der Entwicklung multimedialer Lernsoftware für die Statistikausbildung eine entscheidende Rolle.<sup>63</sup>

**Statistische Software, neue Medien in der Statistikausbildung** Der Einsatz des Rechners in der Statistikausbildung war und ist mit der Frage verbunden, wie dieser die Statistikausbildung verbessern und für welche statistische Arbeiten er eingesetzt werden kann. Der Rechner hilft uns nicht weiter, wenn es z.B. um die Interpretation von Wahrscheinlichkeiten oder Analyseergebnissen geht.

"Computers allow a statistical class to consider the complexities of the real world."<sup>64</sup>

Bei dem Einsatz der neuen Medien soll der Rechner nicht ausschließlich zur Analyse komplexer Datensätze oder Simulation von Realitäten eingesetzt werden. Er übernimmt zusätzlich die Aufgabe, Studierende für die Statistik zu motivieren. Die neuen Medien bieten viele neue Möglichkeiten, das Interesse der Studierenden zu wecken, z.B. benutzerfreundliche Oberflächen, Visualisierungen in Form von Applets, Animationen oder Videofallstudien. Neue Vermittlungsinstrumente werden in Teil III dieser Arbeit genauer betrachtet. An dieser Stelle soll die zentrale Komponente der entwickelten Lernsoftware *Statistik interaktiv komplett* vorgestellt werden: das Statistiklabor.

**Grundlegendes zum Statistiklabor:** Das Statistiklabor ist eine interaktive Arbeitsumgebung zur Modellierung, Bearbeitung und Lösung statistischer Fragestellungen. Die Oberfläche bietet ein Arbeitsblatt, auf dem Daten erzeugt, importiert, bearbeitet und auf verschiedene Weise analysiert und ausgegeben werden können. Als Werkzeuge zur Erfüllung dieser Aufgaben stehen hierbei graphische Laborobjekte zur Verfügung, die zur Übersicht in drei Kategorien eingeteilt werden können:

- Datenquellen und Datenpräsentation zur Erzeugung, Darstellung bzw. zum Import von Daten (Datensatzimport, Zufallszahlen-Generator, Urliste, Datensatz, Matrix), Datenanalyse (Häufigkeitstabelle, Kontingenztafel, Zeitreihen, R-Kalkulator)
- Möglichkeiten zur grafischen oder textuellen Ausgabe von Daten und zum Export in Fremdformate (Datensatzexport, R-Grafik, R-Grafik-Wizard, Texteditor)
- Mit Hilfe des Konnektors (Verbindungspfeil zwischen einzelnen Objekten) können diese verschiedenen Objekte miteinander verbunden werden und die entsprechenden Datenflüsse sowohl logisch festgelegt als auch grafisch dargestellt werden.

---

<sup>61</sup>[Moore1997].

<sup>62</sup>[Moore1997].

<sup>63</sup>Vgl. 10.

<sup>64</sup>[HoaglinMoore1992], Kap. 3 *Computers and modern Statistics*.

Einige dieser Objekte verfügen über eine systeminterne Schnittstelle zu der statistischen Programmiersprache R<sup>65</sup>. Die umfangreiche Funktionalität von R ist damit im Statistiklabor komplett verfügbar.

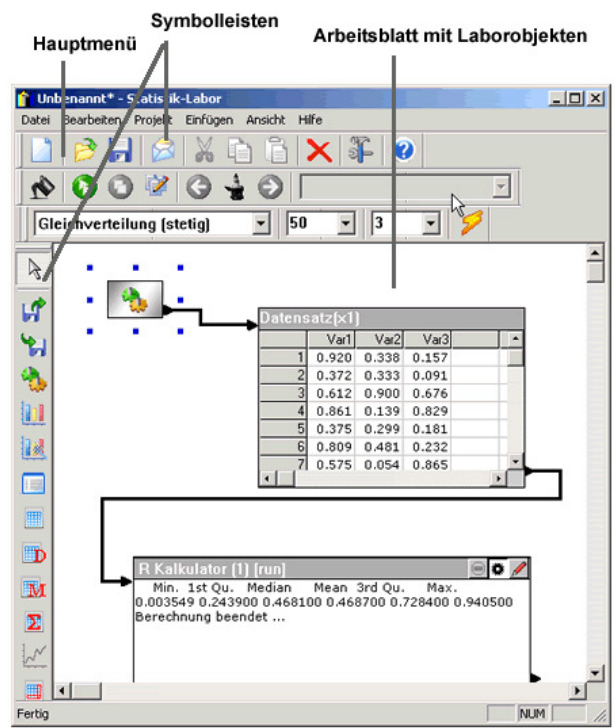


Abbildung 1: Funktionen der Programmoberfläche

Die im Statistiklabor bereitstehenden Objekte können in einer bestimmten Reihenfolge angewendet und durch die Konnektoren (schwarze Pfeile in der obigen Grafik) verknüpft werden: die Reihenfolge entspricht der des statistischen Arbeitens. Analysiert man einen Datensatz im Statistiklabor, so bildet die Urliste oder Datenmatrix den Ausgangspunkt der Datenanalyse. Diese Objekte werden mit dem Objekt *Häufigkeitstabelle* verbunden, um die Daten tabellarisch in übersichtlicher Form darzustellen. Die Objekte *Grafik-Wizard* oder *R-Grafik* werden mit der Häufigkeitstabelle oder einem *R-Kalkulator* verbunden und erzeugen entsprechende Grafiken. Man sieht, dass die Schritte der Datenanalyse in der Reihenfolge durchlaufen werden, wie sie von einem Statistiker gewählt werden. Die Intention besteht darin, den Studierenden durch den Umgang mit den Objekten des Statistiklabors an die Arbeits- und Vorgehensweise des Statistikers heranzuführen.

Das Statistiklabor ermöglicht verschiedene Formen der Vermittlung. Durch die Bereitstellung des *Grafik-Wizards* ist es dem Studierenden möglich, Grafiken zu erstellen, ohne R-Befehle kennen zu müssen. Das Objekt *R-Kalkulator* kann eingesetzt werden, um z.B. Grafiken durch die Verwendung von R-Befehlen zu erzeugen. Je nach Schwerpunktsetzung des Dozenten können die Objekte verschieden intensiv eingesetzt und genutzt werden. Für Dozenten, die eigene Programme oder Funktionen bereitstellen wollen, besteht die Möglichkeit, diese durch sogenannte Bibliotheken einzubinden:

<sup>65</sup><http://www.r-project.org/>

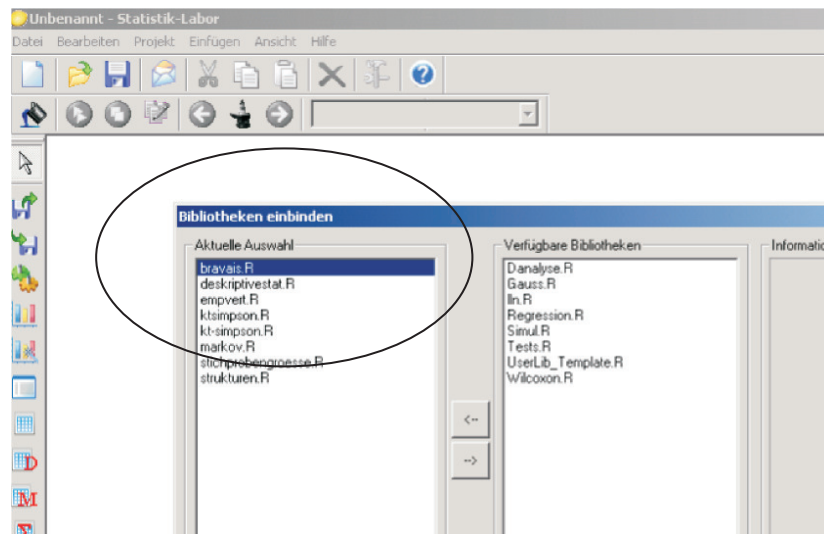


Abbildung 2: Einbinden von Bibliotheken im Statistiklabor

Auch Studierende können eigene Bibliotheken schreiben und im Statistiklabor nutzen. Das Statistiklabor ist somit sowohl im Grund- als auch im Hauptstudium ein hilfreiches Instrument.

Eine weitere Idee des Statistiklabors ist die unkomplizierte Heranführung der Studierenden an komplexes Datenmaterial. Datensätze unterschiedlichen Formats können durch das Objekt *Datensatzimport* eingelesen und analysiert werden. Es wurde versucht, sowohl den Import als auch den Export von Daten so einfach wie möglich zu gestalten, um Barrieren abzubauen und die Studierenden zur Datenanalyse und damit Statistik zu motivieren. Der konkrete Umgang mit den einzelnen Laborobjekten, die Möglichkeiten, die diese en detail bieten, werden dem Leser schrittweise im Verlauf der folgenden Kapitel vorgestellt.

## 2.2 Konstruktivismus, Ziele und neue Vermittlungsmethoden

In Bezug auf den Prozess des Lernens existieren viele entgegengesetzte Meinungen, Theorien und Ansätze. Mit Blick auf das Thema und die gesetzten Schwerpunkte dieser Arbeit erscheint eine historische Betrachtung des Lernprozesses in allen Facetten, die Diskussion klassischer Lerntheorien und deren Eignung für die Statistikausbildung zu weit gefasst. Die Auseinandersetzung mit unterschiedlichen Lerntheorien lässt den Autor zu dem Schluss kommen, dass der Konstruktivismus am besten geeignet erscheint, den identifizierten Defiziten in der Statistikausbildung<sup>66</sup> entgegenwirken zu können. Aus diesem Grund steht der Konstruktivismus als Lerntheorie im Mittelpunkt der folgenden Betrachtungen - unter Beachtung von Glasersfelds:

”Constructivism does not claim to have made earth-shaking inventions in the area of education; it merely claims to provide a solid conceptual basis for some of the things that, until now, inspired teachers had to do without theoretical foundation.”<sup>67</sup>

Ein Problem scheinen alle Lerntheorien trotz ihrer Unterschiede gemeinsam zu haben: die schwierige Messbarkeit des Lernfortschritts/-prozesses.

<sup>66</sup>Vgl. 3.

<sup>67</sup>[Glaesersfeld1987].

Die folgende, stark vereinfachte Darstellung dient der Einordnung des Konstruktivismus. Es existieren zwei Hauptlerntheorien: der Behaviorismus und die sogenannten kognitiven Ansätze, zu denen der Konstruktivismus gezählt wird. Der Behaviorismus kann wiederum aufgespalten werden in die klassische Konditionierung (Stichwort: Pawlow) und die operante Konditionierung (Stichwort: Learning as Problem Solving).

Das Lernen wird von den Vertretern des Behaviorismus durchaus als eine aktive Tätigkeit betrachtet, bestimmt durch ausgelöste Erfahrungen. Das Lernen und der Lernende sind von außen gesteuert. Die Lernenden haben sich verschiedene Verhaltensweisen angeeignet und ändern diese von Zeit zu Zeit, da sie neues Verhalten erwerben. Dieses Verhalten und Wissen stellt im Behaviorismus ein Gut dar, das vom Dozenten (wählen wir den Bereich Universität) an den Studierenden weitergegeben wird. Eine wichtige Grundvoraussetzung für diese Auffassung von Lernen ist die vorhandene Bereitschaft des Studierenden, eine ständige Veränderung seiner Verhaltensweisen von außen zuzulassen. Als die beiden wichtigsten Vertreter des Behaviorismus können Edward Thorndike (1874-1949) und Burrhus Skinner (1904-1990) angesehen werden.

Der Behaviorismus berücksichtigt die Aktivität des Geistes nicht. Deshalb ist er nicht in der Lage, den gesamten Prozess des Lernens zu erfassen und kann nicht alles Lernen erklären. Trotz dieser Schwäche muss positiv hervorgehoben werden, dass der Behaviorismus immerhin die Auffassung vertritt, dass Lernen aktiv stattfindet. Die Aktivität des Lernenden ist ein Aspekt, dem auch beim multimedialen Lernen eine entscheidende Rolle zukommt. Es ist eine der Hauptaufgaben, sowohl bei der Entwicklung multimedialer Lernsoftware als auch bei deren Einsatz darauf zu achten, dass der Studierende durch Interaktionsmöglichkeiten zum Statistikkennen motiviert wird.

Die Vertreter des Konstruktivismus sehen eine ihrer Hauptaufgaben darin, beim Lernenden den Antrieb zum Lernen zu erzeugen. Da der Konstruktivismus zu beschreiben versucht, *wie* Lernende lernten, kann er als eine deskriptive Theorie bezeichnet werden. Piaget, Garfield, Ernst von Glasersfeld zählen zu den wichtigen Vertretern des Konstruktivismus. Diese Arbeit orientiert sich an einigen ihrer Konzepte, Ideen und Aussagen. Die Basis des Konstruktivismus bildet die These, dass Wissen nicht wie beim Behaviorismus von außen gesteuert wird, sondern aus einem ständigen Wechselspiel zwischen dem Lernenden selbst und seiner Umwelt konstruiert wird.

"In the constructivist perspective, knowledge is constructed by the individual through his interactions with his environment."<sup>68</sup>

Piaget bezeichnet den Konstruktivismus als die ständige Interaktion zwischen bereits vorhandenem Wissen und neuen Erfahrungen.<sup>69</sup> Dieser Antrieb zur aktiven Tätigkeit der Wissenskonstruktion muss im Studierenden jedoch erst in Gang gesetzt werden, was wieder Motivation und Bereitschaft erfordert. Die Verbindung zwischen dem Lernenden und seinem Wissen kann wie folgt beschrieben werden: der Lernende verfügt über ein bestimmtes Wissen, auch gewisse Überzeugungen sind in ihm vorhanden. Neues Wissen oder die Erweiterung des existierenden Wissens entsteht durch eine Interaktion zwischen neuen Ideen und Situationen und dem vorhandenen Wissen. Die Vertreter des Konstruktivismus sehen eine der Hauptaufgaben des Lehrers bzw. Dozenten darin, diese Interaktion zu fördern. Dabei bekommt der Lehrer eine neue Rolle. Er muss die Studierenden so zum

---

<sup>68</sup>[Murphy1997].

<sup>69</sup>Vgl. [Tiemann2002], S. 110.

Lernen motivieren, dass sie sich für die Konzepte, bei uns die der statistischen Methoden interessieren.

Der Konstruktivismus hat in den 90er Jahren erstmals Einzug in die Statistikausbildung erhalten. Konstruktivistische Ideen wurden in ersten Versuchen umgesetzt, Inhalte, Pädagogik und Technik dabei zusammen betrachtet. Der diskutierte Wandel in der Statistikausbildung führte dazu, dass nach neuen erfolgreichen, besseren Lernmethoden gesucht wurde, um den existierenden Defiziten zu begegnen. Der Konstruktivismus kann als Beschreibungsrahmen angesehen werden, der sich anbot, während man nach Alternativen suchte. Allerdings gibt der Konstruktivismus keine Anleitung zur praktischen Umsetzung seiner Konzepte und Ideen. Das Lernen wird als Prozess aufgegriffen und neu erklärt, eine Theorie des Lehrens ist aber nicht im Konstruktivismus enthalten. Der Konstruktivismus sagt nicht, wie Lernen und Lehren tatsächlich gestaltet sein soll, verschreibt kein bestimmtes Lern-/Lehrkonzept. Betrachtet man, welche Ideen und Prinzipien beim Konstruktivismus im Vordergrund stehen, so kann abgeleitet werden, dass sich in dem so genannten *problemorientierten Lernen und Lehren* konstruktivistische Elemente am besten widerzuspiegeln scheinen.

Zieht man aus dem Konstruktivismus Konsequenzen für die Defizite in der Statistikausbildung, so muss man vorsichtig sein: z.B. muss eine höhere Aktivität/Interaktivität nicht zwingend zu einem höheren Lernerfolg beim Studierenden führen. Integriert man konstruktivistische Elemente in bestehende Unterrichtsmethoden, so müssen sich auch die Zielsetzungen des Dozenten sowie die Verhaltensweisen der Studierenden entsprechend ändern.

Ein wichtiges Ziel besteht darin, von der bloßen Vermittlung von Faktenwissen wegzukommen, hin zu der Vermittlung von Fähigkeiten, die den Studierenden ein Leben lang von Nutzen sind. Fakten ändern sich in unserer Zeit schnell, veralten, werden durch neue abgelöst. Deshalb ist es wichtig, dem Studierenden neben Fachwissen, die Fähigkeit zu vermitteln, Wissen selbst zu konstruieren, aus neuen Fakten, neuen Informationen, die das Informationszeitalter im Eiltempo bereitstellt. Diese Wissenskonstruktion ist eine wichtige Idee des Konstruktivismus.

”(...), learning emphasizes the process and not the product. How one arrives at a particular answer, and not the retrieval of an objective true solution, is what is more important. Learning is a process of constructing.”<sup>70</sup>

Wir können weitere zu vermittelnde Fähigkeiten benennen: konzeptionelles Verständnis, die Fähigkeit, kritisch zwischen Alternativen abzuwägen, sinnvoll Entscheidungen zu treffen und Ergebnisse numerischer Analysen inhaltlich zu interpretieren.

Es steht nicht zur Debatte, auch die konkreten in unserem Fall statistischen Methoden, die sogenannten *hardskills* zu vermitteln - dies kann als die notwendige Bedingung angesehen werden. Aber man muss weitergehen und den Studierenden die eben genannten Fähigkeiten mit auf den Weg geben, um ihnen wichtige Entscheidungen zu erleichtern, sie zu kritischen Köpfen auszubilden.

**Pragmatischer Konstruktivismus** Aus der Überlegung heraus, traditionelle und konstruktivistische Elemente zu kombinieren, ist der *Pragmatische Konstruktivismus* entstanden, der zunehmend an Bedeutung gewonnen hat. Als wichtiger Vertreter kann Perkins genannt werden. Der pragmatische Konstruktivismus plädiert für folgendes Vorgehen: er

---

<sup>70</sup>[Murphy1997].

nimmt sich vom Konstruktivismus nur das, was er braucht, benutzt diesen wie eine Art Sammlung guter Ideen und Konzepte, die zur Lösung bestimmter Probleme geeignet sind. Im Prinzip entsteht ein Kompromiss aus Konstruktivismus und klassischen Lerntheorien: wo die klassischen Lerntheorien versagen, wird versucht, durch Anwendung des passenden konstruktivistischen Elementes zu einer Lösung zu gelangen.<sup>71</sup> Diesem Ansatz werden wir bei der angestrebten Integration konstruktivistischer und multimedialer Elemente in die Statistikausbildung nachgehen. Die im Folgenden diskutierten Ziele der neuzugestaltenden Statistikausbildung stehen in engem Zusammenhang zum Konstruktivismus.

### Ziele der Statistikausbildung - neue Vermittlungsmethoden

Die Ziele der Statistikausbildung haben sich im Rahmen der Neugestaltung verändert. Wir erinnern uns an die Differenzierung in P- und Mstatistics und das Ziel, Pstatistics stärker in der Ausbildung zu verankern. Mit der Veränderung der Ziele ändern sich auch die Vermittlungsmethoden. Verschiedene Statistiker, Pädagogen, Didaktiker haben neue Vermittlungsmethoden vorgeschlagen und diskutiert, inwiefern sich die neuen Ziele besser erreichen lassen. Es werden die Ausführungen Gals/Garfields vorgestellt und diskutiert, da sie differenziert vorgehen und der Richtung dieser Arbeit entsprechen.

Im Rahmen der Dissertation wird die These aufgestellt, dass Ziele nur dann erreicht werden, wenn die Vermittlung in einem Konzept organisiert ist.

Betrachten wir die von Gal und Garfield aufgestellte Liste<sup>72</sup> von Zielen der Statistikausbildung. Diese werden jeweils erweitert durch Umsetzungsmöglichkeiten in *Statistik interaktiv komplett*:

"Although the depth or breath of these seven goals may differ according to educational level, they describe the main goals for all students who learn basic statistics."<sup>73</sup>

- **Goal 1: Understand the purpose and logic of statistical investigations:** Studierende sollen verstehen, warum statistische Untersuchungen durchgeführt werden und welche statistischen Ideen diesen zugrunde liegen. Ideen wie allgegenwärtige Variabilität in Daten, deren Visualisierung und Benutzung numerischer Zusammenfassungen sollen vermittelt werden. Des Weiteren sollen die Studierenden verstehen, warum Stichproben gezogen werden anstatt die Grundgesamtheit zu untersuchen. Die im Rahmen dieser Arbeit entwickelten multimedialen Lernmodule illustrieren Ideen statistischer Methoden insbesondere durch den Einsatz des Statistiklabors. So wird z.B. die Stichprobenziehung mit Hilfe verschiedener Visualisierungen vermittelt. Ausführlich wird dies in dem Beispiel zur Vermittlung der Varianzzerlegung am Ende dieses Kapitels demonstriert.<sup>74</sup>
- **Goal 2: Understand the process of statistical investigations:** Es wird angestrebt, den Studierenden den Prozess und die Natur des statistischen Arbeitens zu vermitteln. Dies umfasst den gesamten Prozess von der Erhebung der Daten bis hin zur Auswahl geeigneter Methoden. Ziel ist es, jede Phase des statistischen Problemlösungsprozesses zu beleuchten: die Formulierung geeigneter Fragen, Planung der Analyse, Sammlung und Organisation der Daten, die Datenanalyse selbst sowie ihre Interpretation und Präsentation ihrer Ergebnisse.

---

<sup>71</sup>Vgl. [Perkins1999].

<sup>72</sup>Vgl. [Garfield2000], S. 100ff.

<sup>73</sup>[Garfield2000], S. 101.

<sup>74</sup>Vgl. 4.

Dieser Aspekt wird in dem Verfassen statistischer Reports<sup>75</sup> aufgegriffen, das eine zentrale Komponente des im Rahmen dieser Arbeit entwickelten Einsatzkonzeptes darstellt. Außerdem spiegelt sich das statistische Arbeiten in der Gestaltung des Labors und der Verwendungsreihenfolge der Laborobjekte wider.

- **Goal 3: Learn statistical skills (Master procedural skills):** Dieses Ziel der Statistikausbildung betrifft die sogenannten *hardskills*, d.h. das Beherrschen der Umsetzung statistischer Methoden, sowohl manuell als auch am Rechner. Es sollen sowohl die Ideen und Anwendungsfelder der einzelnen Methoden, als auch deren Schwächen und Stärken vermittelt werden. Dies betrifft die Organisation und Durchführung statistischer Untersuchungen, sowie verschiedene Arten der Datenpräsentation. Die zum Ende dieses Kapitels präsentierten Lösungsvorschläge demonstrieren besonders den Aspekt der Umsetzung statistischer Methoden im Statistiklabor. Die Vorteile der neuen Medien, insbesondere des Labors werden dabei betont. Es wird herausgestellt, dass das *wie* des Einsatzes eine entscheidende Rolle spielt.
- **Goal 4: Understand probability and chance (Understand mathematical relationships):** Studierende sollen darin unterstützt werden, ein Verständnis für Konzepte und Modelle zu entwickeln. Sie sollen verstehen, wie z.B. Wahrscheinlichkeitsmodelle helfen können, Entscheidungen unter Unsicherheit zu messen und dadurch besser treffen zu können. Es muss vermittelt werden, wie entsprechende Modelle entwickelt und benutzt werden, um Zufälligkeiten zu simulieren und Daten zu produzieren, mit denen Wahrscheinlichkeiten geschätzt werden können. Das im Rahmen der Arbeit entwickelte Einsatzkonzept<sup>76</sup> sieht einen hohen Einsatz des Statistiklabors besonders bei der Vermittlung stetiger Verteilungsmodelle vor. Außerdem enthält die Lernsoftware *Statistik interaktiv komplett* Theoriekomponenten, die die Modelle erklären, sowie verschiedene Visualisierungen (Applets, Animationen) zum Transport ihrer Ideen.
- **Goal 5: Develop interpretive skills and statistical literacy:** In diesem Ziel steckt der Anspruch, Studierenden alle wichtigen Einzelheiten des Interpretationsprozesses einer statistischen Untersuchung zu vermitteln. Dies beinhaltet die kritische Reflexion der numerischen Ergebnisse, die Fähigkeit, mediale Darstellungen kritisch zu betrachten und statistische Berichte von Kommilitonen kritisch zu hinterfragen. Ein Schwerpunkt der Entwicklung multimedialer Lernmodule und entsprechender Einsatzkonzepte liegt auf dem statistischen Report. In statistischen Reports werden genau die von Gal und Garfield in diesem Goal 5 angesprochenen Aspekte vermittelt: die Interpretation und kritische Reflexion numerischer Ergebnisse steht im Vordergrund eines statistischen Reports.
- **Goal 6: Develop useful statistical dispositions:** Studierende sollen durch die Statistik(ausbildung) ein Gefühl für die Rolle von Zufall entwickeln. Sie sollen verstehen, dass sie dieses Wissen in statistischen Untersuchungen genauso benötigen wie in Businessentscheidungen oder sozialen Entscheidungen, die sie aufgrund unvollständiger Informationen unter Unsicherheit treffen. Sie müssen ein Bewusstsein dafür entwickeln, jedes Argument, das auf Daten und deren Analyse und Interpretation/Repräsentation beruht, kritisch zu hinterfragen.

---

<sup>75</sup>Vgl. 12.2.2.

<sup>76</sup>Vgl. 14.2.



Um dieses Ziel zu erreichen, wurden in der Lernsoftware Fallstudien implementiert. Diese beschreiben Situationen des Lebens, in denen Entscheidungen unter Unsicherheit getroffen werden müssen: es werden Daten zur Entscheidungsfindung herangezogen und entsprechend analysiert.

- **Goal 7: Develop statistical reasoning (Develop the ability to communicate statistically):** Es wird angestrebt, den Studierenden die Fähigkeit zu vermitteln, in statistischer Weise zu kommunizieren. Dazu gehört, statistische Informationen zu verstehen, statistische Ideen zu begründen, Daten zu interpretieren, diese geeignet zu präsentieren, aus statistischen numerischen Zusammenfassungen Informationen zu gewinnen, Informationen über Daten und Zufall sinnvoll zu kombinieren. Die Erreichung dieses Zieles wird durch den Einsatz der Komponente *Statistisches Report* angestrebt.

Die diskutierten Ziele sprechen zwei unterschiedliche Ebenen an: die technische Ebene, auf der die statistischen Methoden erlernt werden sollen (Mstatistics) und die Ebene der Problemlösungskompetenz oder statistischen Kommunikation (Pstatistics).

Zusammenfassend kann zu den Zielen und der Bedeutung der Statistikausbildung Scheaffer zitiert werden:

”Arguably, though, the most important function of statistics education is to prepare people in general to be intelligent consumer(s) of data.”<sup>77</sup>

Die Ziele stellen neue Herausforderungen an die Statistikausbildung. Diesen wird im Folgenden durch den Einsatz der neuen Medien begegnet. Es wird außerdem überprüft, inwiefern die Lernmodule der *Neuen Statistik* den Gedanken gerecht werden, die Vardeman and Morris in ihrem Artikel *Some Advice for Young Statisticians*<sup>78</sup> fordern:

” (...) that statistics can be used to form highly technical and even technically correct support for statements which are in fact not true. We might hope this could happen only when nonstatisticians practice statistics without proper technical understanding of the subject. But statistical lies are by definition immoral uses of statistical arguments, whether technically correct or not, and stem from social pressures that affect statisticians and nonstatisticians alike.”<sup>79</sup>

”Statisticians must be knowledgeable about the system under study. They should not present themselves as competent to analyze data from systems about which they have no substantive understanding. Real data are not ”context-free”. (...) Never forget that you are not the context expert.”<sup>80</sup>

” (...) you cannot afford to think of yourself as a data technician or a hired gun. You must be secure enough to simultaneously separate and prior vested interest (yours or others’) in the outcome from your analysis, and meld together seamlessly everything you know about the subject matter of your investigation with the structure of your statistical work.”<sup>81</sup>

---

<sup>77</sup>[Scheaffer2002], S. 6.

<sup>78</sup>Vgl. [VardemanMorris2003].

<sup>79</sup>[VardemanMorris2003], S. 25.

<sup>80</sup>[VardemanMorris2003], S. 25.

<sup>81</sup>[VardemanMorris2003], S. 26.

Auch in diesen Aussagen/Apellen wird die Notwendigkeit der Kombination von P- und Mstatistics deutlich - vor allem die stärkere Betonung von Pstatistics in der neuen Statistikausbildung.

### **Definition eigener Ziele der Statistikausbildung für diese Arbeit**

Die Auseinandersetzung mit dem Wandel und der Neugestaltung der Statistikausbildung führt zu der Definition von acht Zielen oder besser Zieleigenschaften der Studierenden. Die Vermittlung dieser statistischen Kompetenzen stellt das Ziel des Einsatzes von *Statistik interaktiv komplett* in der Statistikausbildung dar. Ob diese erreicht wurden, mit Hilfe welcher Komponenten und mit welchen Formen des Einsatzes, wird die Evaluation überprüfen.<sup>82</sup> Die erwünschten Fähigkeiten stellen somit einen erwünschten Sollzustand der Studierenden (als Output der Statistikausbildung) dar:

- Verständnis für statistische Methoden - Methodenverständnis
- Verständnis für gezielten (und richtigen) Einsatz statistischer Methoden zur Problemlösung - Anwendungsverständnis
- Verständnis für Zusammenhänge und Abhängigkeiten zwischen den Methoden
- Verständnis für mathematische Konzepte und schrittweise Beweisführung
- Zur Dokumentation des statistischen Arbeitens: Fähigkeit, statistische Reports zu verfassen
- Souveräner Umgang mit neuen Medien und Verständnis für die notwendige Zusammengehörigkeit von "Computational" und "Statistics"
- Fähigkeit, organisiert und selektiv zu lernen, Prioritäten beim Lernen zu setzen
- Erster Überblick über Anwendungsfelder, Relevanz und Rolle der Statistik in der Realität

Im Rahmen der Evaluationsstudie<sup>83</sup> wird diskutiert, wie diese Zieleigenschaften erreicht werden könnten. Dabei werden sowohl die Prinzipien des Konstruktivismus, als auch die Möglichkeiten der neuen Medien genutzt.

Im folgenden Abschnitt wird ein weiteres Problem der Statistik/Statistiker betrachtet: das Imageproblem.

---

<sup>82</sup>Vgl. 14.2.

<sup>83</sup>Vgl. 14.2.

### 2.3 Statistical Bridges

Richard L. Scheaffler<sup>84</sup> beschreibt in seinem 2002 im Journal of the American Statistical Association erschienenen Artikel *Statistical Bridges* ein immer noch existentes Problem der Statistiker: "In the information age of today, statistics is essential but statisticians are not."<sup>85</sup> Scheaffler weist durch historische Zitate bekannter Statistiker darauf hin, dass die Probleme der Statistiker nicht ganz neu sind und dass sich Statistiker zu sehr unterschiedlichen Zeitpunkten bereits geäußert haben:<sup>86</sup>

**Ellenberg 2000:** statistics receives more recognition than statisticians

**Moore 1998:** our future as a discipline depends on our taking this broad view

**Kettenring 1997:** the need to reconstruct the image of statistics

**Bradley 1982:** the future of statistics as a discipline was in some danger and urged us to mitigate this trend by improving our image, making sure that statisticians get recognized for their contributions of importance to society, and working together for the mutual benefit of all.

**Waker 1944:** Ways must be found to make the general nature of statistical thinking better understood by the average intelligent citizen and particularly by persons (...) charged with policy formation (...) If the public fails to value and support statistical research, we much take some of the blame ourselves in that we have failed to insist that the nature of statistical thinking is an appropriate topic for inclusion in a liberal education.

Scheaffler betont, dass das Informationszeitalter beides bewirken kann - den jetzigen Zustand verschlechtern oder aber das Ansehen der Statistiker verbessern.

"The demand for statistics may be growing faster than the demand for statisticians, but even the latter is growing faster than the supply. There is hope for the future, but bringing hope to reality depends on increasing our numbers and widening our vision, while simultaneously improving the public perception of statistics."<sup>87</sup>

Scheaffler sieht die Statistik als eine Insel zwischen vielen Inseln an. Dabei stellt jede Insel eine Disziplin dar. Scheaffler betont die Notwendigkeit der regen Kommunikation zwischen den Disziplinen, plädiert für Interdisziplinarität. „We must develop our island without being isolated. We may dwell on other islands for a while and may even build cottages there, but we should not live there permanently. There should always be a tie to our home island of statistics, which nourishes us for further adventures.“<sup>88</sup>

Die durch die Brücken hergestellten Verbindungen zu anderen Disziplinen können als Instrument verstanden werden, die Rolle und Stärke der Statistiker und der Statistik hervorzuheben. Scheaffler unternimmt den Versuch, solche Brücken zu schlagen.<sup>89</sup> Im Folgenden wird Brücke 3 exemplarisch betrachtet:

---

<sup>84</sup>Professor Emeritus, Department of Statistics, University of Florida, Gainesville

<sup>85</sup>[Scheaffler2002], S. 1.

<sup>86</sup>Vgl. [Scheaffler2002], S. 1.

<sup>87</sup>[Scheaffler2002], S. 1.

<sup>88</sup>[Scheaffler2002], S. 2.

<sup>89</sup>Die Beschreibung der Brücken ist angelehnt an Scheafflers Ausführungen. Es werden allerdings nur die Aspekte aufgegriffen, die die Autorin im Rahmen dieser Arbeit für interessant erachtet. Der interessierte Leser kann ausführliche Erläuterungen nachlesen bei [Scheaffler2002].

**Brücke 3: College Curriculum - Academic concentrations for a few** Diese Brücke betrifft das sogenannte *Undergraduate Program* der Statistikausbildung amerikanischer Universitäten. Es soll eine Verbindung zu denen hergestellt werden, die speziellere Statistikkenntnisse erwerben wollen. Natürlich sollen die Studierenden statistische, mathematische und *computational* Kenntnisse erwerben, doch sollte der Anspruch an die Lehre auf dieser Stufe noch etwas höher geschraubt werden. Um diesem gerecht werden zu können, wurde vor drei Jahren eine *ASA*<sup>90</sup> *undergraduate statistics education initiative* gegründet. Diese entwickelte folgende Punkte, gedacht als Leitlinie der zukünftigen neu gestalteten Statistikausbildung:<sup>91</sup>

- Betonung realer Daten und authentischer Anwendungen
- Präsentation der Daten in einem Kontext, der den Studierenden anspricht
- Erfahrungen mit "statistical computing"
- Stärkere Synthese von Theorie-Anwendung
- Verstärkte Entwicklung statistischer Kommunikationsfähigkeiten

In dieser Arbeit werden sowohl die multimedialen Komponenten der Lernmodule, als auch die speziell entwickelten Lernmodule, als auch das für diese entwickelte Einsatzkonzept aufgegriffen. Diese sind darauf ausgerichtet, jedem dieser Aspekte gerecht zu werden.<sup>92</sup>

Durch die Existenz verschiedener Brücken wird deutlich, dass die Statistik zahlreiche Bereiche betrifft: unterschiedliche Ausbildungsstufen, die Praxis, die Daten benutzt, die Wissenschaft sowie auch die Öffentlichkeit und ihre Vertreter. In jedem dieser Bereiche werden Statistik und somit Statistiker benötigt. Diese müssen gut ausgebildet sein, um ihren Aufgaben gewachsen zu sein und Fehlentscheidungen zu verhindern. Wir schließen uns Scheaffer an, der darauf hinweist, dass Statistik eine zu entscheidende Rolle spielt, um eine Ausbildung zu dulden, bei der jeder vor sich hinlehrt, die Statistik nicht in den Zusammenhang mit anderen Disziplinen setzt und sich nur selbst amüsiert. - "In short, statistics must be an island without being insular."<sup>93</sup>

Es sei wiederholt, dass die Existenz und Veröffentlichung eines solchen Artikels in einer renommierten statistischen Zeitschrift als Leitartikel im Jahr 2002 deutlich machen sollte, dass es sich bei der Neugestaltung der Statistikausbildung um ein aktuelles Thema handelt.

Scheaffer verdeutlicht die Notwendigkeit eines differenzierten Blickes auf die Problematik. Dieser wird durch die Benennung verschiedener Brücken zwar dargestellt, allerdings fehlt das *was*. Welche Art von Statistik soll konkret über diese Brücken transportiert werden? In multimedialen Vorschlägen werden verschiedene Dimensionen der Statistik vermittelt bzw. transportiert.<sup>94</sup>

**Zusammenfassung und Überleitung** Der Wandel der Statistikausbildung und die existierenden Defizite wurden überblickartig beschrieben. Ziele der neu zu gestaltenden Statistikausbildung wurden abgeleitet. Die in den bisherigen Ausführungen angesprochenen

<sup>90</sup>American Statistical Association

<sup>91</sup>Vgl. [Scheaffer2002], S. 4.

<sup>92</sup>Die konkrete Umsetzung wird in den Teilen III und IV der vorliegenden Arbeit diskutiert.

<sup>93</sup>[Scheaffer2002], S. 6.

<sup>94</sup>Vgl. 4.

Defizite werden im Folgenden klassifiziert und in Verbindung zueinander gestellt. Zentrale Defizite werden anhand statistischer Beispiele demonstriert. Es werden ausgearbeitete Vorschläge unterbreitet, diesen mit Hilfe der neuen Medien entgegenzuwirken.

### 3 Defizitklassifikation

Die vorliegende Arbeit geht von einer Defizitklassifikation aus. Die in der Statistikausbildung existenten Defizite werden zusammenstellt und klassifiziert. Zusammenhänge zwischen den Defizitgruppen werden aufgezeigt. Die einzelnen Defizite werden durch Beispiele illustriert, wobei diese oft mehr als ein Defizit der Klassifikation ansprechen. Jedes Beispiel enthält die folgenden Aspekte: es wird beschrieben, worin das Defizit bzw. die Defizite bestehen. Außerdem wird erläutert, wie bereits versucht wurde, dem durch die traditionelle und/oder rechnergestützte Statistikausbildung entgegenzuwirken. Als Konsequenz der Existenz der Defizite wird ein Lösungsvorschlag auf Basis der neuen Medien entworfen. Dieser wird mit Hilfe der im Rahmen dieser Arbeit mitentwickelten multimedialen Statistikernsoftware *Statistik interaktiv komplett* umgesetzt. Es handelt sich dabei nicht zwingend um völlig neue Ansätze. Vielmehr geht es darum, die Ansätze der bisherigen Ausbildung aufzugreifen, zu verbessern und weiterzuentwickeln.

**Zusammenstellung existenter Defizite** Die voranstehenden Ausführungen haben Schwierigkeiten der Statistikausbildung beschrieben und teilweise diskutiert. Diese werden im Folgenden durch zusätzliche Defizite erweitert und zusammengefasst dargestellt. Es ergeben sich die folgenden Defizite:<sup>95</sup>

- **A:** Die Disziplin Statistik wird allgemein nicht wichtig genommen, ihre Möglichkeiten werden unterschätzt. Es wird nicht erkannt, wie komplex sie ist, dass sie sich aus drei wichtigen Teilen zusammensetzt: den statistischen Methoden, dem statistischen Problemlösen und *Computational Statistics*. Außerdem wird die statistische Dimension realer Probleme nicht erkannt.
- **B:** Es fehlt an Interaktivität/Aktivität der Studierenden, die für das Erlernen der Statistik notwendig ist. Die Lehre fördert die Interaktivität/Aktivität der Studierenden ungenügend, folgt dem Prinzip *Learning as a constructive activity*<sup>96</sup> zu wenig.
- **C:** Der interdisziplinäre Charakter von Statistik wird nicht vermittelt. Brücken zu anderen Disziplinen werden nicht geschlagen und somit die Relevanz der Statistik nicht klar. Auch die Relevanz der Statistik für Entscheidungsprozesse wird unterschätzt.
- **D:** Die Dominanz der methodischen, mathematischen Seite der Statistik stellt eine Schwierigkeit dar: Studierende verstehen Statistik oft als Sammlung von Methoden,

---

<sup>95</sup>Bemerkung: Betrachten wir die folgenden zwei Defizite:

- Der Statistik wird an Hochschulen nicht genug Platz eingeräumt. Die Zeit, Statistik zu vermitteln und zu lernen, ist zu knapp bemessen.
- Vorurteile der Studierenden gegenüber der Statistik (mathematisch, schwer, langweilig, Zahlenpyramiden, Statistiken).

Beide Defizite existieren, beide betreffen die Statistik(ausbildung). Sie sind nicht in die aufgestellte Defizitklassifikation aufgenommen, da man ihnen im Rahmen unseres Vorhabens nicht direkt entgegenwirken kann.

<sup>96</sup>[Moore1997].

stellen die statistische Analyse in den Vordergrund. Es mangelt an konzeptionellem Verständnis, Reflexion, Dokumentation und Interpretation statistischer Analyseergebnisse (Statistischer Report). Es findet eine Anwendung statistischer Methoden statt, ohne zu wissen, warum diese in bestimmten Situationen angewendet werden.

- **E:** Die Studierenden verstehen die Ursache für die Entstehung statistischer Paradoxe nicht, sind deshalb nicht vorbereitet auf Situationen, in denen diese auftreten könnten.
- **F:** Die mathematische Seite statistischer Methoden bereitet vielen Studierenden Schwierigkeiten. Sie verstehen die Mathematik nicht, was ein fehlendes Methodenverständnis zur Folge hat.
- **G:** Computational Statistics (Kombination von Statistik und Rechner) scheint wichtig für bestimmte statistische Konzepte zu sein, stellt aber gerade für den Anfänger eine Hürde dar. Die Brücke zwischen angewandter Statistik und dem Einsatz des Rechners wird nicht geschlagen.
- **H:** Die Zusammenhänge einzelner statistischer Methoden/Begriffe/Konzepte werden oft nicht vermittelt. Es kommt bei dem Studierenden zu einer Kenntnis verschiedener Methoden. Er kann diese aber in keinen sinnvollen Zusammenhang bringen.
- **I:** Es besteht die folgende Schwierigkeit: Mangelnde Motivation der Studierenden für die Statistik führt zu mangelnder Bereitschaft, statistische Konzepte zu erlernen.
- **J:** Die mangelnde Bereitschaft, statistische Konzepte verstehen zu wollen, kann Folgeschäden in der Gesellschaft hervorrufen, verursacht durch Fehlentscheidungen.

**Bildung von Defizitgruppen** Es fällt auf, dass einige Defizite die Ursache für andere darstellen. Es kann zwischen verschiedenen Typen von Defiziten unterschieden werden. Eine Unterscheidung zwischen übergeordneten bzw. Ursachedefiziten, Defiziten und Folgedefiziten erscheint sinnvoll. Die Ursachen für Defizite können sowohl beim Studierenden als auch beim Dozenten liegen. Ist der Dozent der Verursacher, so kann dies häufig auf mangelnde Fähigkeiten und mangelndes Engagement zurückgeführt werden. Beim Studierenden liegt es oft an der mangelnden Bereitschaft und Motivation zum Statistikkernen. Bei der Klassifizierung ist es entscheidend festzulegen, nach welchen Kriterien diese durchgeführt werden soll. Zwei Möglichkeiten zu klassifizieren wurden bereits angesprochen. Die eine besteht darin, nach dem Kriterium *Verursacher* zu klassifizieren. Ein zweites Klassifizierungskriterium könnte mit *Zusammenhänge der Defizite* bezeichnet werden. Da uns weniger daran gelegen ist, Verursacher, sondern Defizite generell zu identifizieren, um ihnen mit neuen Maßnahmen entgegenwirken zu können, wird die folgende Klassifizierung von Defiziten nach dem Kriterium *Zusammenhänge* erfolgen. Das Ziel der Bildung geeigneter Defizitgruppen besteht darin, eine erste Ordnung in die existierenden Defizite zu bringen. In der Abbildung 3 ist die Defizitklassifikation für die Statistikausbildung dargestellt, welche die folgenden Gruppen enthält:

- **Übergeordnetes Defizit: A**  
Die Unterschätzung der Disziplin Statistik kann häufig darauf zurückgeführt werden, dass nicht erkannt wird, dass sie sich aus mehr als statistischen Methoden zusammensetzt. Diese Unterschätzung sehen wir im Rahmen dieser Arbeit als übergeordnetes Defizit an, da jedes der anderen Defizite mit diesem in Verbindung steht. Ein

Beispiel: Die Dominanz der methodischen Vermittlung der Statistik und der Vernachlässigung des Verfassens statistischer Reports kann daraus resultieren, dass in der Lehre verkannt wird, wie wichtig das Verständnis statistischer Methoden ist. Es findet eine Unterschätzung statt: Statistik wird als reine Sammlung von Methoden betrachtet und vermittelt. Dies kann auch daraus resultieren, dass der Statistik zu wenig Zeit in den einzelnen Disziplinen eingeräumt wird, der Dozent keine Zeit für die Vermittlung statistischen Denkens hat. Wieder findet eine Unterschätzung der Disziplin statt, hier durch die Hochschulleitung bzw. die in den einzelnen Fakultäten für die Zeiteinteilung Verantwortlichen.

- **Ursachedefizite:** B, C, D, E, F, G, H

Jedes dieser Defizite kann zur Folge haben, dass die Studierenden wenig Bereitschaft zeigen, statistische Konzepte zu erlernen, die Defizite I und J entstehen als Konsequenz. Oft wirken mehrere Defizite gleichzeitig, da sie nicht überschneidungsfrei sind. Ein Beispiel zu den Defiziten B (Interaktivität) und G (Computational Statistics): es bestehen Interdependenzen zwischen der Interaktivität der Studierenden und dem Einsatz des Rechners in der Statistikausbildung. Wie wir in einzelnen Lösungsvorschlägen später sehen werden, kann der gezielte Einsatz des *Statistiklabors*, die Interaktivität des Studierenden fördern. Er wird durch den Rechner zum experimentellen, aktiven Lernen angeregt. Fehlender Einsatz des Rechners könnte fehlende Interaktivität zur Konsequenz haben. Beide zusammen würden die Folgedefizite I und J hervorrufen.

- **Folgedefizit:** I, J

Beide Defizite können als Folgen der anderen Defizite angesehen werden, wobei das Defizit I direkt folgt, das Defizit J indirekt: Die mangelnde Bereitschaft der Studierenden, Statistik zu erlernen, kann dazu führen, dass sie statistische Konzepte nur halb oder gar nicht verstehen und im Beruf Fehlentscheidungen treffen, die zu gesellschaftlichen Konsequenzen führen.

**Strukturierung und Defizitklassifikation** An die Klassifizierung der Defizite, ihre Zuordnung zu einzelnen Defizitgruppen schließt sich deren Strukturierung an: es wird eine Struktur gebildet, die die Zusammenhänge zwischen den Defizitgruppen darstellt. Das Ergebnis dieser strukturierten Klassifizierung der Defizite in der Statistikausbildung ist eine mögliche Defizitklassifikation:<sup>97</sup>

---

<sup>97</sup>Durch die Pfeile sind die Ursache-Wirkungsbeziehungen zwischen den Defizitgruppen gekennzeichnet.

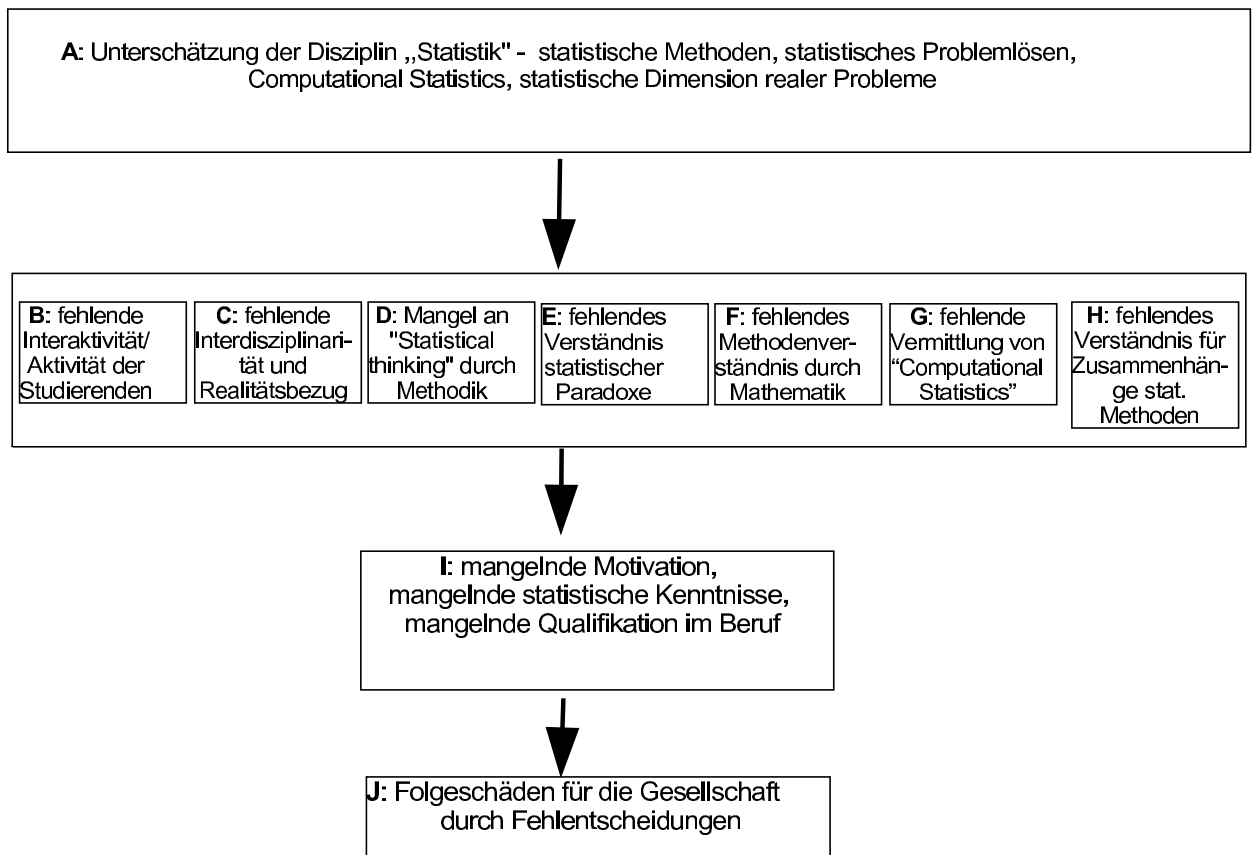


Abbildung 3: Defizitklassifikation für die Statistikausbildung

Sicher ist dies nicht die einzige Möglichkeit, die Defizite in der Statistikausbildung zu betrachten. Die Gestalt einer Defizitklassifikation ändert sich in Abhängigkeit von den Klassifizierungskriterien, die zugrunde gelegt werden. Der Vorteil dieser Struktur liegt in der Klarheit: es ist sofort ersichtlich, welches Defizit zu welchen Folgen führt und auf welcher Ebene es angesiedelt ist. Die Defizitklassifikation bildet die Grundlage für die weitere Argumentation dieser Arbeit. Einzelne Defizite werden exemplarisch verdeutlicht. Es werden konkrete Vorschläge unterbreitet und umgesetzt, ihnen entgegenzuwirken.



## 4 Multimediale Lösungsvorschläge

In diesem Kapitel werden vier Umsetzungen vorgestellt, mit deren Hilfe den Defiziten der Statistikausbildung unter Nutzung der neuen Medien entgegengewirkt werden kann. Ausgewählte statistische Methoden, bei denen oben klassifizierte Defizite auftreten, werden durch Einsatz der statistischen Lernsoftware *Statistik interaktiv komplett* neu aufbereitet. Es entstehen vier multimediale Lösungsvorschläge, die die Defizite beheben sollen. Diese Lösungsvorschläge sprechen verschiedene Dimensionen der Statistik an, spannen somit einen Raum der Einsatzmöglichkeiten multimedialer Lernsoftware in der Statistikausbildung auf. Dieser wird überblickartig skizziert:

- Varianzzerlegung: ein **statistisches Konzept** wird exemplarisch erläutert und durch konstruktivistisch gestaltete multimediale Komponenten vermittelt.
- Simpson-Paradox: die **Interpretation und der Umgang mit einem statistischen Konstrukt**, der Kontingenztabelle steht im Mittelpunkt der Diskussion. Auch hier stellt der Konstruktivismus ein wichtiges Element bei der Gestaltung der Laborseiten dar.
- Strukturen in Daten: Der Umfang realer Daten und die Schwierigkeit der Statistik, Strukturen in diesen zu erkennen, wird beleuchtet. **Transformationsprozesse** werden eingesetzt, um Strukturen zu erkennen. Der Rechner wird als Hilfsmittel zur Durchführung verschiedener Transformationsprozesse und grafischen Darstellungen eingesetzt.
- Problem der geeigneten Stichprobengröße: ein klassisches **Problem der Statistik, die Stichprobenziehung**, wird demonstriert. Wann ist eine Stichprobe repräsentativ, d.h. lässt Rückschlüsse auf die interessierende Grundgesamtheit zu? Dieses Problem sowie die Problematik der Beschaffenheit der Stichprobe wird exemplarisch anhand realer Daten diskutiert.

Die vier multimedialen Umsetzungen/Lösungsvorschläge können daher in beliebiger Reihenfolge gelesen werden. Sie besitzen Demonstrationscharakter. Unterschiedliche Dimensionen der Statistik können multimedial angesprochen und in Bezug auf die Ausbildung verbessert werden. Die Qualitäten und Möglichkeiten des Statistiklabors werden herausgestellt. Die Lösungsvorschläge demonstrieren zudem die verschiedenen Arten der technischen Umsetzung statistischer Methoden im Statistiklabor. Es werden Laborobjekte eingesetzt, um statistische Probleme zu lösen. Einige Statistiklaborseiten benutzen voreingestellte Objekte, die auf Mausklick hin statistische Grafiken erzeugen, andere binden komplizierte selbst programmierte Funktionen/Programme (in Form von Bibliotheken) ein, um statistische Vorgänge abzubilden. Auch hier wird ein Raum von Möglichkeiten aufgespannt und dem Leser demonstriert, der Raum der technischen Möglichkeiten des Statistiklabors in Bezug auf die Umsetzung statistischer Methoden.

#### 4.1 Methodenverständnis - illustriert an der Varianzzerlegung

Defizit F<sup>98</sup> besagt, dass Studierende statistische Methoden oft nicht verstehen, da sie mit deren mathematischen Aspekten Schwierigkeiten haben. Dieses Defizit wird anhand der Varianzzerlegung exemplarisch diskutiert. Die multimediale Lernsoftware *Statistik interaktiv komplett* wird eingesetzt. Es wird ein konkreter Vorschlag ausgearbeitet, einzelnen Defiziten entgegenzuwirken.<sup>99</sup>

Ein Teil des Lösungsvorschlags ist in sogenannten Theoriekomponenten enthalten. Der Beweis der Varianzzerlegung wird in aufeinander aufbauende Schritte zerlegt, die Schritte finden in einzelnen Theoriekomponenten statt. Diese Zerlegung folgt einer Grundidee des Konstruktivismus: die Studierenden konstruieren Wissen, indem sie bereits erlerntes und neues Wissen schrittweise vernetzen.

**Idee der Varianzzerlegung, Defizit, Vorgehensweise** Die Idee der Varianzzerlegung besteht darin, die Gesamtstreuung eines zusammengesetzten Datensatzes, der aus  $k$  Datensätzen besteht, durch

- die Streuung innerhalb der Gruppen und
- die Streuung zwischen den Gruppen

zu erklären. Die folgenden Betrachtungen erläutern die Varianzzerlegung auf Basis des Streuungsmaßes MQA (Mittlere quadratische Abweichung). Für diese sieht die Varianzzerlegung wie folgt aus:

$$d^2 = \frac{n_i}{n} d_i^2 + \sum_{i=1}^k \frac{n_i}{n} (\bar{x}_i - \bar{x})^2$$

Der erste Term der rechten Seite der Gleichung beschreibt die mittlere quadratische Abweichung innerhalb der Gruppen, der zweite die zwischen den Gruppen.

Die Idee der Varianzzerlegung besteht darin, einen Teil der Streuung eines Datensatzes durch eine Zerlegung auf Merkmale zurückzuführen. Liegt bei einem Datensatz eine Struktur vor, die einen Hinweis auf die Existenz unterschiedlicher Gruppen liefert, können mit Hilfe der Varianzzerlegung Teile der Streuung auf diese Gruppen zurückgeführt werden. Die Erklärung der Streuung kann z.B. anhand des Merkmals *Geschlecht* durchgeführt werden. Die gesamte Streuung des ursprünglichen Datensatzes wird in die Streuung innerhalb der Gruppe der Männer, innerhalb der Gruppe der Frauen und die Streuung zwischen diesen beiden Gruppen zerlegt. Durch Anwendung der Varianzzerlegung kann ein Teil der Gesamtstreuung auf das Geschlecht bzw. die Gruppen zurückgeführt werden.

In diesem Zusammenhang tritt ein Defizit auf, das sich aus den beiden folgenden Komponenten zusammensetzt:

- Die Studierenden verstehen die Idee der Varianzzerlegung nicht, erkennen die Vorteile, die durch diese entstehen können, somit nicht;

---

<sup>98</sup>Vgl. 3.

<sup>99</sup>Bemerkung: Ein Teil des Lösungsvorschlags besteht darin, die Idee der Varianzzerlegung grafisch zu illustrieren. Das bedeutet nicht, dass die Visualisierung als Königsweg zur Mathematik angesehen wird. Es handelt sich lediglich um *ein* Instrument, welches eingesetzt wird, um die Idee der Varianzzerlegung zu transportieren.

- Die Studierenden haben Schwierigkeiten, die mathematische Formel, ihre Zusammensetzung aus verschiedenen Komponenten zu verstehen. Sie besitzen häufig ein schwach ausgeprägtes Formelverständnis, können deshalb dem Beweis der Varianzzerlegung nicht folgen.

Schlägt man den Bogen zu den obigen Ausführungen, so lautet die Konsequenz: die Studierenden haben bezüglich der Varianzzerlegung weder die Seite der Pstatistics noch die der Mstatistics verstanden.

In der traditionellen Statistikausbildung gibt es Vertreter, die dafür plädieren, den Beweis der Varianzzerlegung nicht zu vermitteln, sondern nur deren Idee. Diese Arbeit schließt sich den Vertretern an, die für die Vermittlung mathematisch statistischer Beweise plädieren, da die Mathematik als ein wichtiger Bestandteil der Statistik betrachtet wird:

”Mathematics uses precise language and precise statements. Those statements are justified by logically proofs. Mathematics unifies examples and generalizes them to a broader theory. I argue that statistics can and should use these approaches to become more relevant to the problems that it tackles, not less. Moreover, where a corpus of mathematical theory is available which solves the statistical problem at hand, it should be used. Finally, some statistical communication could benefit from the clear labelling used by mathematicians.”<sup>100</sup>

Die durch die Grafiken vermittelte Idee der Varianzzerlegung wird außerdem mathematisch erklärt. Zudem wird der Zusammenhang der Quadratsumme, der mittleren quadratischen Abweichung und der Stichprobenvarianz formal beschrieben. Die sich schrittweise ergebende Formel der Varianzzerlegung ist gut geeignet, um die statistische Idee und die mathematische Präzision zu demonstrieren.

Im Folgenden wird die Varianzzerlegung in aufeinander aufbauenden Schritten bewiesen und sowohl an einem realitätsnahen als auch einem fiktiven Beispiel erläutert. Es wird anhand eines Beispiels aus der Wirtschaft vermittelt, in welchen Situationen die Varianzzerlegung angewendet werden kann und zu einem direkten Informationsgewinn führt. Es sei angemerkt, dass die Varianzzerlegung nicht in allen Fällen, in denen sie angewendet werden kann, zu Informationsgewinnen führt.

**Vorgehensweise und Umsetzung** In Theoriekomponenten der multimedialen Lernsoftware *Statistik interaktiv komplett* wird der Beweis geführt und umgesetzt. Die Komponenten sind miteinander verlinkt. Sie sind außerdem mit Statistikkaborseiten verbunden, die die Idee der Varianzzerlegung graphisch illustrieren. Außerdem könnten z.B. Verlinkungen der Theoriekomponenten zu Beispielkomponenten und Glossareinträgen erfolgen. Alle multimedialen Komponenten zusammengefasst bilden eine zusammenhängende Einheit zur Vermittlung der Idee und des Beweises der Varianzzerlegung.

---

<sup>100</sup>[Bailey1998], S. 261.

Die folgende Übersicht stellt die Überlegungen grafisch dar:

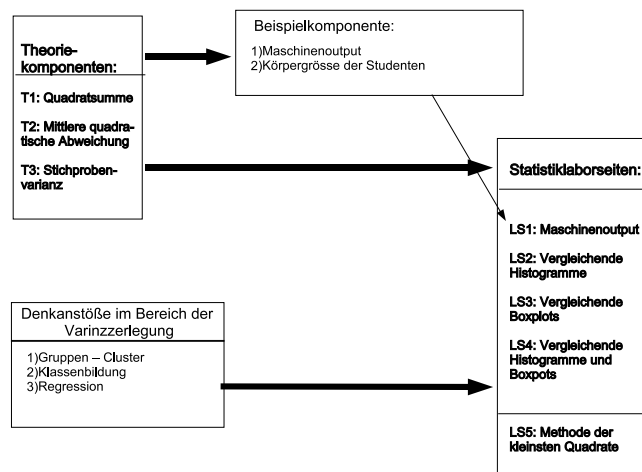


Abbildung 4: Übersicht zur Vorgehensweise

Im Folgenden betrachten wir die schrittweise Beweisführung innerhalb der miteinander verlinkten Theoriekomponenten.

Die Theoriekomponenten sind durch Verlinkungen auch mit den entsprechenden Statistiklaborseiten verbunden. Auf drei Statistiklaborseiten wird die Idee der Varianzzerlegung grafisch dargestellt. Theoriekomponenten und Statistiklaborseiten wirken in Kombination miteinander.

Dem Studierenden wird die mathematische Seite der Varianzzerlegung innerhalb der Theoriekomponenten vermittelt, die Statistiklaborseiten verdeutlichen die dahinter stehende Idee. Der kombinierte Einsatz ist konzipiert für eine Vorlesung oder Übung, in der der Dozent die verschiedenen Komponenten präsentiert.<sup>101</sup> Auch hier folgt die Gestaltung konstruktivistischen Prinzipien: der Studierende kann sein Wissen selbst konstruieren, indem er die durch verschiedene Medien vermittelten Inhalte zueinander in Verbindung setzt.

Die in der Grafik als *Denkanstöße* bezeichneten Aspekte werden in einem abschließenden Unterkapitel kurz diskutiert. Sie können als eine Art Ausblick für den Dozenten betrachtet werden, der *Statistik interaktiv komplett* sowie den folgenden Vorschlag einsetzen und erweitern möchte.

**Die Idee der Varianzzerlegung - praktisch** Warum ist die Varianzzerlegung überhaupt interessant? Ein Beispiel, das im wirtschaftlichen Kontext angesiedelt ist, soll die Studierenden dazu motivieren, die Varianzzerlegung zu verstehen.<sup>102</sup>

<sup>101</sup>Erfahrungen der Evaluationsstudie haben gezeigt, dass die aktive Nutzung des Angebots ohne den Dozenten, d.h. ohne *guided learning/teaching*, nicht erfolgt.

<sup>102</sup>Der Text könnte in einer Beispielkomponente von *Statistik interaktiv komplett* erscheinen.

Ein Studierender könnte in fünf Jahren Leiter der Personalabteilung sein. Das ihn beschäftigende Unternehmen produziert Schrauben auf vier baugleichen Maschinen. Die Schrauben unterliegen Vorgaben, da sie in die extern produzierten Muttern passen müssen. Ein hoher Anteil der produzierten Schrauben wird den Vorgaben zur Zeit nicht gerecht. Die Untersuchung aller Maschinen zusammen hat ergeben, dass im Mittel den Vorgaben gemäß produziert wird. Die Untersuchung der Gesamtstreuung aller Maschinen zeigt, dass ein zu hoher Anteil der Produktion aus dem vorgegebenen Intervall (1314.5 - 1315.5 Mikrometer Durchmesser) herausfällt. Die Geschäftsleitung plädiert für den Kauf von vier neuen Maschinen, was aus Kostengründen Entlassungen hervorrufen würde.

Der ehemalige Studierende setzt sich in seiner neuen Position das Ziel, diese Entlassungen zu verhindern und wendet dafür seine Statistikenkenntnisse an: er beschafft sich Zusatzinformationen über die einzelnen Maschinen. Eine Untersuchung jeder Maschine für sich ergibt, dass jede Maschine im Mittel die Bestimmungen erfüllt. Er zieht den Schluss, dass der hohe produzierte Ausschuss auf die maschinenspezifische Streuung zurückzuführen ist und bestimmt für jede einzelne Maschine die jeweilige Produktionsvarianz. Das Ergebnis zeigt, dass lediglich zwei der vier Maschinen ersetzt werden müssen, da die Streuungen der anderen beiden Maschinen im vorgegebenen Intervall liegen. Durch die Zusatzinformation ist es also möglich, einen Teil der Gesamtstreuung durch Streuungen einzelner Maschinen zu erklären. Diese Erkenntnis spart die Anschaffungskosten für zwei Maschinen und kann Entlassungen verhindern.

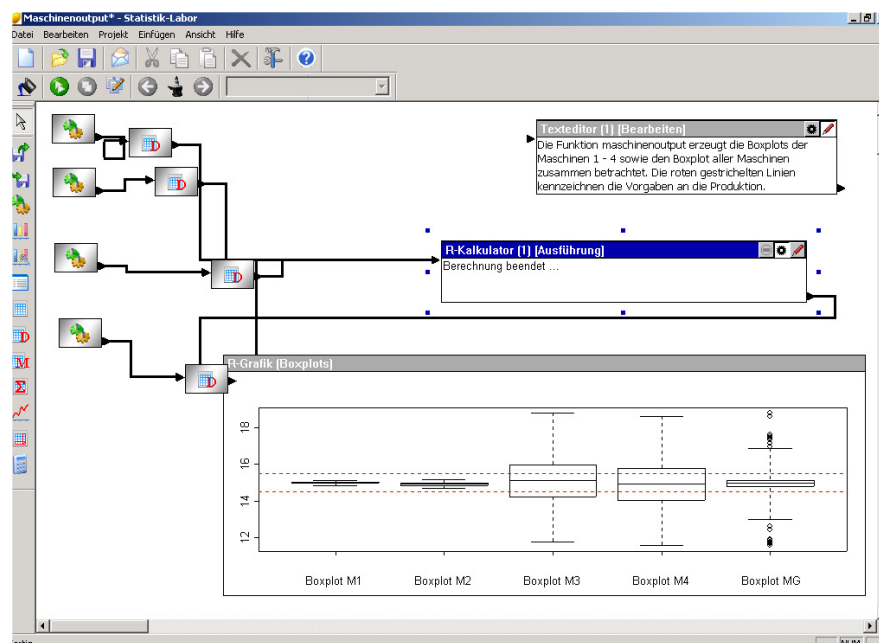


Abbildung 5: Varianzzerlegung am Beispiel im Statistikkolabor umgesetzt

Die in den einzelnen Maschinen enthaltene Streuung wird durch Boxplots dargestellt. Die Studierenden erkennen, wie sich die Gesamtstreuung (Boxplot MG) zusammensetzt und sehen, dass lediglich die Maschinen drei und vier (Boxplot M3 und M4) für die starke Gesamtstreuung verantwortlich sind. Der Einstieg über das Beispiel verdeutlicht, warum die Varianzzerlegung nicht nur ein statistisches Konstrukt ist, sondern dass sie auch wirtschaftswissenschaftliche Entscheidungsprozesse positiv vorantreiben kann.

**Der Beweis - in Schritte zerlegt, in Theoriekomponenten dargestellt** Der Beweis der Varianzzerlegung wird in drei Teile zerlegt. Der erste Teil enthält den Beweis über die Zerlegung der Quadratsumme. Der zweite Teil baut auf diesem auf und beweist die Varianzzerlegung über die Zerlegung der mittleren quadratischen Abweichung. Dabei werden die Erkenntnisse über die Zerlegung der Quadratsumme angewendet. Im dritten Teil wird analog über die Stichprobenvarianz argumentiert. Jeder Teil wird durch eine eigene Theoriekomponente erfasst. Der Studierende kann lernen, in Analogien zu denken, Wissen zu vernetzen. Das Wissen über die Zerlegung der Quadratsumme und der mittleren quadratischen Abweichung kann er benutzen, um die Varianzzerlegung auf Basis der Stichprobenvarianz zu verstehen.

Die Beweisführung wird auf Basis mathematischer Sätze geführt. Diese werden während der einzelnen Schritte angegeben. Der Studierende muss die Fähigkeit besitzen, logische Schlüsse selbst zu ziehen und mathematische Sätze auf Konstrukte anzuwenden bzw. zu übertragen. Diese Vorgehensweise soll den Studierenden zum aktiven Erlernen statistischer/mathematischer Zusammenhänge anregen. Es wird ihm kein Beweis vollständig erklärt, sondern er muss Wissenskonstruktion betreiben, indem er Bekanntes auf Neues überträgt und neue Zusammenhänge herstellt. Der Konstruktivismus spricht vom sogenannten *active learner*.

Die Theoriekomponente soll exemplarisch demonstrieren, wie die mathematischen Hintergründe und die Beweisführung gestaltet sein könnten. Insbesondere soll verdeutlicht werden, was mit der obigen Aussage *Der Studierende muss die Fähigkeit besitzen, logische Schlüsse selbst zu ziehen und mathematische Sätze auf Konstrukte anzuwenden bzw. zu übertragen*. konkret gemeint ist.

Theoriekomponente 1 - Microsoft Internet Explorer

Datei Bearbeiten Ansicht Favoriten Extras ?

Zurück Suchen Favoriten Medien

Adresse C:\Dokumente und Einstellungen\apabst\Lokale Einstellungen\Temp\wbt-Varianzzerlegung2\html\1a0c.html

THEMA: MATHEMATISCHE SEITE DER STATISTIK

Die Varianzzerlegung

Theoriekomponente 1

vor

f. THEORIEKOMPONENTE 1

In dieser Theoriekomponente wird die Zerlegung der Summe der Quadrate hergeleitet. Es wird auf Basis Ihnen bekannter mathematischer Sätze argumentiert. Dies setzt bei Ihnen die Fähigkeit voraus, diese auf komplexere Strukturen anwenden zu können und logische Schlüsse zu ziehen.

Den Ausgangspunkt bilden die beiden folgenden Formeln:

- die Berechnung des arithmetischen Mittels:

$$\bar{x} = \sum_{r=1}^n x_r$$

- die Berechnung der Summe der quadrierten Abweichungen der einzelnen Beobachtungen vom arithmetischen Mittel:

$$\sum_{r=1}^n (x_r - \bar{x})^2$$

Die einzelnen Beweisschritte enthalten Angaben zu dem jeweils relevanten mathematischen Hintergrund:

Für die mittlere quadratische Abweichung der einzelnen Gruppen gilt:

$$d_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (\bar{x}_{ij} - \bar{x})^2$$

$$\sum_{r=1}^n (x_r - \bar{x})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2 \text{ mathematischer Hintergrund: Kommutativgesetz}$$

$$\sum_{r=1}^n (x_r - \bar{x})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i + \bar{x}_i - \bar{x})^2 \text{ mathematischer Hintergrund: Nulladdition}$$

An dieser Stelle ist eine Nebenrechnung notwendig:

$$\begin{aligned} ((x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x}))^2 &= (x_{ij} - \bar{x}_i)^2 + 2(x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x}) + (\bar{x}_i - \bar{x})^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} 2(x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x}) + \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 \end{aligned}$$

mathematischer Hintergrund: 2. Binomische Formel

Um einen Teil dieser Formel zu erklären, bietet sich folgende Nebenrechnung an:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} 2(x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x}) = 2 \sum_{i=1}^k (\bar{x}_i - \bar{x}) \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)$$

mathematischer Hintergrund: Assoziativgesetz und Kommutativgesetz

$$\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i) = 0$$


---

statistischer Hintergrund: Schwerpunkteigenschaft des arithmetischen Mittels

Die Anwendung der Schwerpunkteigenschaft führt zu folgender Beziehung:

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 &= \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 \\ \Rightarrow \sum_{r=1}^n (\bar{x}_r - \bar{x})^2 &= \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 \end{aligned}$$

Wir sehen die folgende Zerlegung: Die Quadratsumme des Gesamtdatensatzes setzt sich zusammen aus der Quadratsumme innerhalb der Gruppen und der Quadratsumme zwischen den Gruppen.

Abbildung 6: Theoriekomponente 1 umgesetzt in *Statistik interaktiv komplett*

Die zweite Theoriekomponente interpretiert die Zerlegung der Quadratsumme für den Fall der mittleren quadratischen Abweichung (MQA). Es findet keine neue Beweisführung statt. Der Studierende soll die Inhalte der ersten Theoriekomponente aufgreifen und das Wissen für die MQA selbst konstruieren (Konstruktivistisches Prinzip):

The screenshot shows a web browser window titled 'Theoriekomponente 2 - Microsoft Internet Explorer'. The address bar shows a local file path. The page content includes a navigation bar with 'zurück' and 'vor' buttons, a section header 'THEORIKOMPONENTE 2', and a paragraph explaining the derivation of the decomposition of the sum of squares for the mean square deviation. The mathematical derivation is as follows:

$$d^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_i \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \frac{1}{n} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

$$= \frac{1}{n} \sum_{i=1}^k n_i d_i^2 + \frac{1}{n} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

Below the equations, a paragraph states: 'Es gilt also: die gesamte mittlere quadratische Abweichung setzt sich zusammen aus der gewichteten mittleren quadratischen Abweichung innerhalb der Gruppen und der gewichteten mittleren quadratischen Abweichung zwischen den Gruppen.'

Abbildung 7: Theoriekomponente 2 umgesetzt in *Statistik interaktiv komplett*



THEORIEKOMPONENTE 3

Auf den voranstehenden Theoriekomponenten aufbauend wird in dieser Theoriekomponente die Zerlegung der Stichprobenvarianz  $s^2$  gezeigt:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^k \frac{n_i - 1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \frac{1}{n} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 + \frac{1}{n-1} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^k (n_i - 1) s_i^2 + \frac{1}{n-1} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

Die Gesamtvarianz setzt sich also zusammen aus der gewichteten Varianz innerhalb der einzelnen Gruppen und der gewichteten Varianz zwischen den Gruppen.

Abbildung 8: Theoriekomponente 3 umgesetzt in *Statistik interaktiv komplett*

**Konkrete Umsetzung im Statistiklabor** Die folgenden Statistiklaborseiten stellen Möglichkeiten zur grafischen Darstellung der Varianzzerlegung dar. Es werden Daten verwendet, die an den Studierenden erhoben wurden. Die Schuhgröße der Studierenden wurde ausgewählt, da die Daten gut geeignet sind, um den Sinn der Varianzzerlegung zu zeigen. Der bimodale Charakter der Daten legt die Anwendung der Varianzzerlegung nahe. Eine weitere Herangehensweise bestünde darin, die Varianz des Datensatzes zu berechnen und zu überlegen, wie der hohe Wert erklärt werden könnte. Die Idee ist bei beiden Ansätzen dieselbe: ich erkenne etwas grafisch oder numerisch Auffälliges. Um dies zu erklären, wende ich die statistische Methode der Varianzzerlegung an.<sup>103</sup>

1. Laborseite a: Der Datensatz *Schuhgröße* wird unterteilt in *Schuhgröße-Männer* und *Schuhgröße-Frauen*. Der Vergleich der Stabdiagramme zeigt, dass die Varianzen der beiden Datensätze unterschiedlich sind. Zunächst wird das Stabdiagramm des gesamten Datensatzes erstellt. Der bimodale Charakter der Daten bietet die Anwendung der Varianzzerlegung an.

<sup>103</sup>Diese Vorgehensweise demonstriert dem Studierenden Statistisches Denken bzw. Problemlösung.

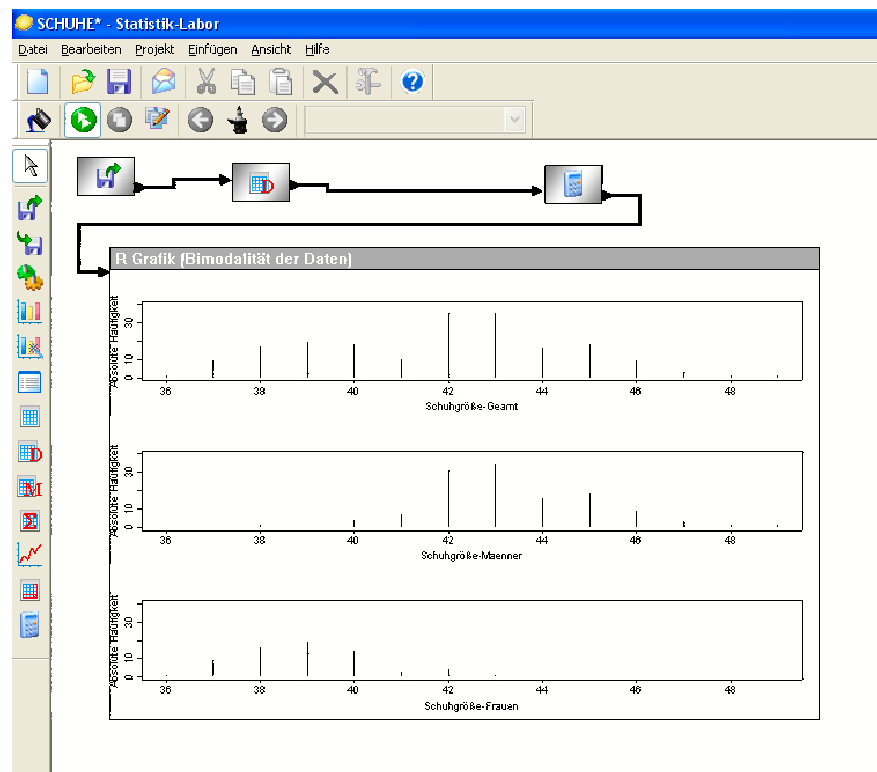


Abbildung 9: Vergleich der Stabdiagramme

Das erste Stabdiagramm zeigt dem Studierenden, dass sich der zusammengefasste Datensatz aus zwei Gruppen zusammensetzen könnte. Ein Einsatzkonzept könnte die Diskussion der Laborseite in der Vorlesung vorsehen. Der Dozent könnte die Studierenden zu der folgenden Erkenntnis leiten: die Betrachtung der Stabdiagramme der Gruppen *Männer* und *Frauen* erklärt die Zusammensetzung der Varianz des zusammengefassten Datensatzes. Diese setzt sich aus der Varianz innerhalb und zwischen den beiden Gruppen zusammen. Außerdem wird er darauf hingewiesen, vergleichende Stabdiagramme heranzuziehen, um eine einheitliche Skala beim Vergleich zugrunde zu legen.

Um die Idee der Varianzzerlegung zu vertiefen, kann folgende Grafik in der Vorlesung diskutiert werden. Dabei sind die Merkmalsausprägungen der Frauen mit *f*, die der Männer mit *m* bezeichnet:

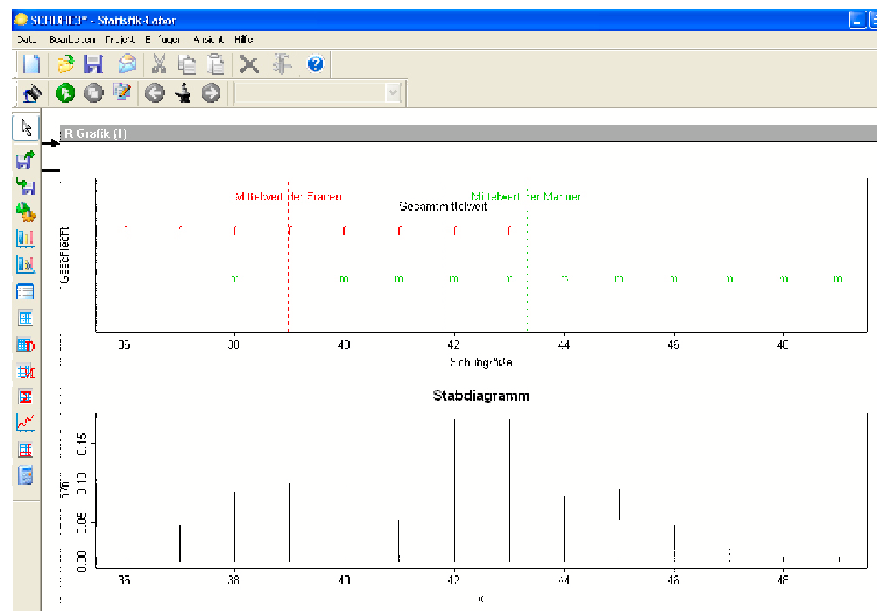


Abbildung 10: Idee der Varianzzerlegung

Die demonstrierten Laborseiten stellen einige Möglichkeiten dar, das Statistikkabor einzusetzen, um die Idee der Varianzzerlegung grafisch zu transportieren. Der Dozent hat allerdings unbegrenzte Möglichkeiten, die dargestellten Laborseiten um weitere zu ergänzen. Der Einsatz ist abhängig vom jeweiligen Einsatzkonzept. Ein konstruktivistisch geprägtes Einsatzkonzept strebt den *active learner* an, der sich Wissen selbst konstruiert. An dieser Stelle würde das bedeuten, dass er durch mehrere Laborseiten und die Vernetzung des so erlernten Wissens den Sinn der Varianzzerlegung erlernt. Die Erfahrungen zeigen aber, dass eine bloße Bereitstellung von Materialien nicht effizient ist. Es empfiehlt sich deshalb, die Theoriekomponenten und Statistikkaborseiten in einer Vorlesung zur Varianzzerlegung einzusetzen und zur Nachbereitung bereitzustellen. Auch im Tutorium können einzelne Komponenten, z.B. die Beispielkomponente in Verbindung mit den Statistikkaborseiten, vertiefend eingesetzt werden. *Guided learning* findet statt. Auch das *guided learning* ist im Konstruktivismus begründet. Die Wissensvernetzung findet weiterhin durch den Studierenden statt. Er wird lediglich durch den Dozenten angeleitet.

**Denkanstöße im Bereich der Varianzzerlegung** Bei der Varianzzerlegung sollten die im Folgenden diskutierten Aspekte je nach Datenlage berücksichtigt werden. Diese sind sowohl für den Dozenten als auch für den Studierenden geeignet: Beiden werden Denkanstöße mit auf den Weg gegeben. Es liegt im Ermessen des jeweiligen Dozenten, diese ausführlicher zu gestalten und zu vermitteln.

- Zuordnung zu Gruppen: Wir sind bisher davon ausgegangen, dass die Gruppen bekannt sind: Männer-Frauen, Maschinen 1-4. Das Problem kann allerdings auch darin bestehen, dass nicht bekannt ist, zu welcher Gruppe ein Objekt gehört, d.h. ein Datensatz mit  $n$  Objekten liegt vor, eine Gruppenstruktur ist nicht bekannt, aber vorhanden. Bevor also mit der Varianzzerlegung begonnen werden kann, steht man vor dem Problem, die Gruppen zu identifizieren. Es gibt verschiedene Verfahren, Gruppenstrukturen in Datensätzen zu identifizieren. Zum Beispiel kann die Diskriminanzanalyse als multivariates Verfahren herangezogen werden. Betrachten wir ein klassisches Beispiel aus dem Bereich der Wirtschaftswissenschaften, Kreditwürdigkeit eines Kunden: „Bei der Vergabe des Kredites ist nicht bekannt, ob der Kunde die Verpflichtungen einhalten wird. Man kennt aber eine Reihe von Merkmalen wie das Alter, das Einkommen und das Vermögen. Auf Basis dieser Informationen ordnet man die Person entweder der Gruppe der Kunden zu, die kreditwürdig sind, oder der Gruppe der Kunden, die nicht kreditwürdig sind.“<sup>104</sup>
- Gruppierung: Die Ergebnisse der Varianzzerlegung sind stark abhängig von der gewählten Gruppierung. Wenn ein ungeeignetes Kriterium zur Klassenbildung verwendet wird, sind die Ergebnisse der Varianzanalyse nicht hilfreich. Werden ungeeignete Gruppen gebildet, so wird das zur Folge haben, dass die Varianz innerhalb der Gruppen sehr groß ist. Wichtig ist hierbei das Verhältnis zwischen der Gesamtvarianz und der Varianz innerhalb der Gruppen des Datensatzes. Die Gruppenvarianz für sich betrachtet lässt absolut gesehen keine Rückschlüsse auf die Größe der Streuung zu. Diese kann auch durch z.B. große numerische Werte der Merkmalsausprägungen hervorgerufen werden. Es besteht die Möglichkeit, dies im Rahmen der einfaktoriellen Varianzanalyse zu untersuchen: mit Hilfe des F-Tests wird die mittlere Streuung innerhalb der Gruppen mit der mittleren Streuung zwischen den Gruppen verglichen.
- Regression: Es besteht die Möglichkeit, die Studierenden im Rahmen der Varianzzerlegung auf den Zusammenhang zu dem linearen Regressionsmodell aufmerksam zu machen. Das am häufigsten angewendete Verfahren, eine Regressionsgerade durch eine Punktwolke zu legen, ist die *Methode der kleinsten Quadrate*. Das Ziel besteht darin, eine Funktion zu bestimmen, die die durch die Datenpaare entstehende Punktwolke so gut wie möglich approximiert. Die Funktion soll sich den Daten anpassen, die wesentliche Struktur des Zusammenhangs zweier Merkmale widerspiegeln. Die Grundidee der Methode der kleinsten Quadrate zur Schätzung der Regressionsgeraden besteht in der Bildung der Quadratsumme. Da diese auch der Varianzzerlegung zugrundeliegt, kann die Methode der kleinsten Quadrate im Zusammenhang mit der Varianzzerlegung vermittelt werden.
- In Verbindung mit der Varianzzerlegung bietet es sich an, sowohl den Studierenden als auch möglichen Dozenten die Bedeutung der Methodenauswahl vor Augen zu führen. Werden Daten durch klassische statistische Methoden zu stark reduziert, können bestimmte Charakteristika der Daten nicht mehr erkannt werden. Dies kann anhand der Boxplots der Schuhgröße der Studierenden im Statistiklabor gezeigt werden.

---

<sup>104</sup>[Handl2002], S. 315.

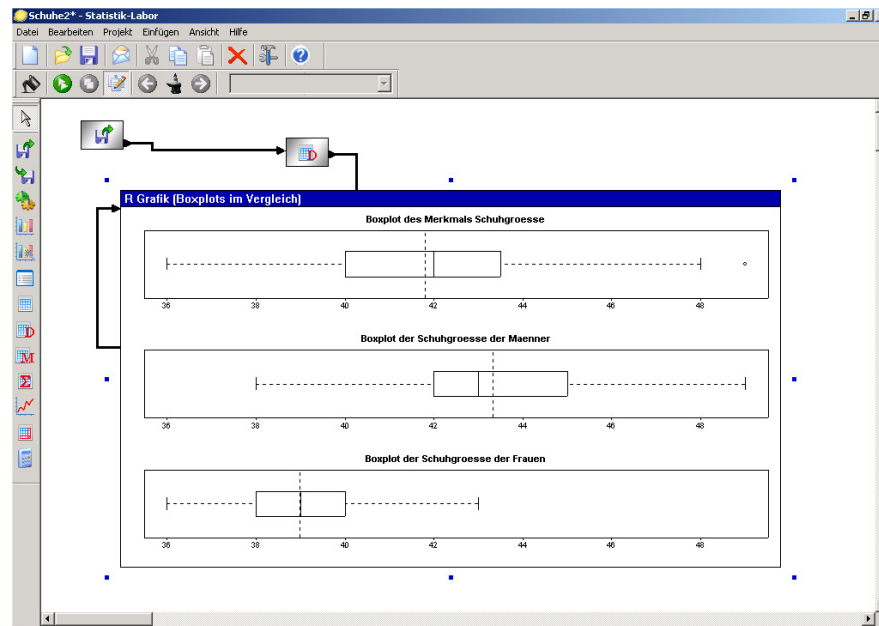


Abbildung 11: Nichtidentifikation der Bimodalität

Betrachtet man den Boxplot des gesamten Datensatzes, so ist keine Multimodalität der Daten erkennbar. Man würde nicht auf die Idee kommen, die Grundgesamtheit in Männer und Frauen aufzuteilen. Dies würde aber verhindern, mit Hilfe der Varianzzerlegung zu dem Ergebnis zu kommen: *Frauen haben im Mittel kleinere Füße als Männer.*

Die bestehenden Streuungs- und Lageunterschiede zwischen den beiden Gruppen werden erst dann deutlich, wenn man die Boxplots für die beiden Gruppen einzeln bildet. Zu diesem Schritt würde der Boxplot des gesamten Datensatzes allerdings nicht motivieren.

Fazit: Das Stabdiagramm eignet sich deutlich besser, die Varianzzerlegung zu vermitteln. Um den Studierenden Zusammenhänge, Vor- und Nachteile statistischer Methoden zu vermitteln, ist es allerdings hilfreich, auch anhand der eben vorgestellten Laborseite zu argumentieren. Diese verdeutlicht die Bedeutung der Methodenwahl.

**Bezug zur Defizitklassifikation** Betrachten wir die aufgestellte Defizitklassifikation<sup>105</sup> und identifizieren die Defizite, denen in diesem Lösungsvorschlag entgegengewirkt werden soll:

- Wie schon in der Einleitung dieses Lösungsvorschlags erwähnt, soll primär dem **Defizit F** (Studierende verstehen statistische Methoden oft nicht, da sie mit deren mathematischen Aspekten Schwierigkeiten haben) entgegengewirkt werden. Der Beweis der Varianzzerlegung wird zerlegt und somit schrittweise erklärt. In Verbindung mit Statistiklaborseiten, die parallel zum Beweis die Idee der Varianzzerlegung illustrieren, soll der Studierende die Methode verstehen und nicht durch die Mathematik abgeschreckt werden. Das Beispiel, das die Varianzzerlegung mathematisch in einem realen Kontext beweist, hat die Aufgabe, die Brücke zwischen Mathematik, Statistik und Realität zu schlagen und das Methodenverständnis des Studierenden zu fördern.
- Dem in **Defizit C** fehlenden Realitätsbezug der Statistikausbildung wird insofern entgegengewirkt, als reale Daten (Körpergröße der Studierenden) zur Demonstration und Erklärung der Varianzzerlegung herangezogen werden. Die Studierenden können die Idee der Varianzzerlegung anhand des bimodalen Charakters ihrer eigenen Daten erkennen. Des Weiteren soll das wirtschaftswissenschaftliche Beispiel (Maschinenoutput) die Brücke zwischen der Statistik und dem späteren Berufsleben der Studierenden schlagen. Indem sie erkennen, dass die Statistik dort eine Rolle spielt, sollen sie zu ihrem Erlernen motiviert werden.
- Durch den Einsatz des Statistiklabors wird der Rechner in den Vordergrund gerückt, somit wird dem **Defizit G** (fehlende Vermittlung von "Computational Statistics") entgegengewirkt. Die Studierenden erkennen die Möglichkeiten, Daten im Statistiklabor bzw. durch den Rechner zu untersuchen und grafisch zu veranschaulichen. Sie erkennen die Vorteile von Computational Statistics in Bezug auf die Charakterisierung von Daten.
- Durch die Statistiklaborseiten soll das statistische Denken der Studierenden, der kritische Blick auf Daten geschult werden. Sie erkennen z.B. einen bimodalen Charakter, überlegen, wie dieser zustande gekommen sein könnte, und untersuchen ihn mit Hilfe geeigneter statistischer Methoden. Die Methodik ist Teil des Denkprozesses der Studierenden, rückt aber nicht vor das statistische Denken wie oft in der traditionellen Statistikausbildung. Es wird dem **Defizit D** (Mangel an statistischem Denken durch Methodik) entgegengewirkt.

Indem den angesprochenen Ursachedefiziten entgegengewirkt wird, werden auch die aus diesen resultierenden Folgedefizite positiv bekämpft.<sup>106</sup>

---

<sup>105</sup> Vgl. 3.

<sup>106</sup> Vgl. 3, Defizite I und J.

## 4.2 Simpson-Paradox – Kombination traditioneller Lehre und neuer Medien

**Ziele und angesprochene Defizite** Beginnen wir mit einem Zitat von Wardrop, der hervorhebt, wie wichtig es ist, statistische Paradoxe zu verstehen:

”Schoolchildren routinely learn to identify optical illusions. It is arguably as important that the general public learn to identify statistical illusions.”<sup>107</sup>

Statistische Paradoxe stellen eine Hürde beim Erlernen und Vermitteln der Statistik dar. Im täglichen Leben kann fehlendes Verständnis statistischer Paradoxe zu Fehleinschätzungen oder Fehlentscheidungen führen. In der traditionellen Ausbildung wurde versucht, dem Defizit „Verständnisschwierigkeiten bezüglich statistischer Paradoxe“ entgegenzuwirken, indem Paradoxe anhand realer Beispiele vermittelt wurden. Dies führte zwar dazu, dass sich die Studierenden stärker für die Statistik interessierten, erleichterte es ihnen allerdings nicht, statistische Paradoxe wirklich zu verstehen. Diesem Defizit soll mit Hilfe der Neuen Medien entgegengewirkt werden. Exemplarisch wird dies für das Simpson-Paradox demonstriert. Es werden reale Beispiele gegeben, bei denen das Simpson-Paradox auftritt. Anschließend wird an einem dieser Beispiele erklärt, wie und wann das Simpson-Paradox zustande kommt.

Im folgenden Schritt kommen die neuen Medien zum Einsatz. Die Methoden der traditionellen und multimedialen Statistikausbildung werden kombiniert. Die Ideen der traditionellen werden aufgegriffen und durch die multimediale Statistikausbildung verbessert, indem Interaktivität und Experimente ermöglicht werden. Experimente in dem Sinne, dass der Studierende das Simpson-Paradox durch Variation von Daten selbst hervorrufen bzw. beseitigen kann.

Für eines der Beispiele wird eine Statistikkorbseite erstellt, die das Simpson-Paradox zum einen demonstriert, zum anderen Interaktivität zulässt. Das Ziel besteht darin, den Studierenden durch eigenes Experimentieren an das Simpson-Paradox heranzuführen: als aktiv Lernender sammelt er im Statistikkorb die Erfahrung, in welchen Fällen das Simpson-Paradox auftritt. Der entwickelte Lösungsvorschlag soll verdeutlichen, dass die Interaktivität, die durch die neuen Medien ermöglicht wird, zu einem besseren Verständnis statistischer Paradoxe führen kann. Diese Gestaltung des Lösungsvorschlags beruht auf dem oben beschriebenen konstruktivistischen Prinzip, das von Moore als *Learning as a constructive activity* bezeichnet wurde.

**Simpson-Paradox - Defizit der traditionellen Statistikausbildung** Sowohl Blyth (1972) als auch Haunsperger (1991) fassen in Worte, worum es sich bei dem Simpson-Paradox handelt:

”Simpson’s paradox is the designation for a surprising situation that may occur when two populations are compared with respect to the incidence of some attribute: If the populations are separated in parallel into a set of descriptive categories, the population with higher overall incidence may yet exhibit a lower incidence within each such category.” (Blyth 1972)<sup>108</sup>

---

<sup>107</sup>[Wardrop1995].

<sup>108</sup>[Wagner1982], S.46.

"Simpson's paradox exhibits seemingly deviant behavior where the data generated in independent experiments support a common decision, but the aggregated data support a different outcome."<sup>109</sup>

"All statisticians know of the dangers of ignoring a covariate that is correlated to an outcome variable and an explanatory one. Simple but convincing examples based on real data are, however, in rather short supply. (Simpson 1951)"<sup>110</sup>

**Historischer Rückblick:** Die Schwierigkeit, statistische Paradoxe zu erklären, wurde auch von der traditionellen Statistikausbildung erkannt. Um das Simpson-Paradox realitätsnah zu erklären, wurde es anhand von realen Beispielen vermittelt. Ein Blick in den *American Statistician* zeigt, dass nach realen Beispielen für das Simpson-Paradox gesucht wurde. Der im Jahr 1982 erschienene Artikel "Simpson's Paradox in Real Life" diskutiert in diesem Zusammenhang anhand zweier Beispiele (Bereich: Einkommensteuerraten), wie leicht das Simpson-Paradox auftreten kann. Häufig tritt in der Literatur auch ein Beispiel aus dem Jahr 1910 auf, bei dem Angaben über Tuberkulose Tote in New York und Richmond, U.S.A. verglichen werden.

Argumentiert man, dass ein solches Beispiel die Studierenden nicht interessiert, so kann die traditionelle Statistikausbildung auf einen Artikel aus dem Jahr 1995 verweisen. In "Simpson's paradox and the hot hand in basketball" erklären Wardrop/Robert dieses sehr anschaulich in Bezug auf Basketball. Sport scheint zumindest einen Großteil der Studierenden anzusprechen. Auch in "Ignoring a covariate: An example of Simpson's paradox" wird 1996 im Teacher's Corner darüber berichtet, den Studierenden das Paradox im Zusammenhang mit Rauchverhalten zu vermitteln.

Trotz aller Bemühungen blieb das Defizit "Verständnisschwierigkeiten bezüglich statistischer Paradoxe" weiter bestehen. Die Studierenden konnten zwar motiviert werden, sich mit dem Simpson-Paradox auseinanderzusetzen, verstanden aber nie wirklich, wie es konkret zustandekam. Dies kann zurückgeführt werden auf die reine Darstellung des Simpson-Paradoxes: zwei Kontingenztabelle werden aggregiert und es wird gezeigt, dass die bedingte relative Häufigkeiten der einzelnen Kontingenztabelle ungleich der aggregierten Kontingenztabelle sind. Ein Defizit besteht unserer Meinung nach darin, dass in der traditionellen Statistikausbildung nur Situationen vorgestellt werden, in denen das Simpson-Paradox auftritt. Hier knüpfen die neuen Medien und der im Folgenden demonstrierte Lösungsvorschlag an:

- Das Auftreten des Simpson-Paradoxes,
- das Nicht-Auftreten des Simpson-Paradoxes
- und der Übergang zwischen beiden Zuständen

können durch die Studierenden experimentell im Statistiklabor herbeigeführt werden. Die konkrete Umsetzung wird dem Leser im folgenden Abschnitt gezeigt.

**Simpson-Paradox - Interaktivität im Statistiklabor** Das Ziel besteht darin, das Verständnis des Studierenden für das Simpson-Paradox durch Interaktivität zu fördern. Es wird eine Statistiklaborseite erstellt, die diese Interaktivität ermöglicht, den Studierenden

<sup>109</sup>[Haunsperger1991], S.252.

<sup>110</sup>[Vanderpump1996], S.340.



selbst experimentieren lässt. Auf der Statistiklaborseite werden drei Kontingenztabelle erzeugt, die dritte stellt die Aggregation der beiden ersten dar. Dies wird anhand des folgenden Zahlenbeispiels demonstriert:

Zur Bekämpfung einer Krankheit werden Forschungsstudien durchgeführt, in deren Rahmen zwei Medikamente zur Bekämpfung der Krankheit entwickelt worden sind und nun auf ihre Wirksamkeit hin untersucht werden sollen. Beide Medikamente wurden in zwei Kliniken eingesetzt und erprobt. Die Resultate liegen in Form der folgenden Kontingenztabelle vor:

		A	B	Summe
<b>Klinik I:</b>	Erfolg	60	100	160
	kein Erfolg	20	50	70
	Summe	80	150	230

		A	B	Summe
<b>Klinik II:</b>	Erfolg	40	10	50
	kein Erfolg	80	30	110
	Summe	120	40	160

Die Berechnung der Erfolgsquoten der Medikamente in den Kliniken erfolgt über die bedingten relativen Häufigkeiten und ergibt, dass in beiden Medikament A erfolgreicher ist als Medikament B:

$$\text{Klinik I: } P(\text{Erfolg}|A) = \frac{60}{80} = 0.75 \text{ und } P(\text{Erfolg}|B) = \frac{100}{150} = 0.66667$$

$$\text{Klinik II: } P(\text{Erfolg}|A) = \frac{40}{120} = 0.3333 \text{ und } P(\text{Erfolg}|B) = \frac{10}{40} = 0.25$$

Die Marketingabteilung schlägt vor, die zusammengefassten Ergebnisse in eine Werbeaktion einzubringen. Zu diesem Zweck werden die beiden Kontingenztabelle der Kliniken durch Addition der Zellen zusammengefasst. Es ergibt sich die folgende aggregierte Kontingenztabelle:

	A	B	Summe
Erfolg	100	110	210
kein Erfolg	100	80	180
Summe	200	190	390

Die Berechnung der Erfolgsquoten der Medikamente in den zusammen betrachteten Kliniken erfolgt wieder über die bedingten relativen Häufigkeiten und ergibt, dass Medikament B erfolgreicher ist als Medikament A:

$$P(\text{Erfolg}|A) = \frac{100}{200} = 0.50 \text{ und } P(\text{Erfolg}|B) = \frac{110}{190} = 0.5789$$

Im Wesentlichen liegt die Begründung für das Paradox im Auftreten der folgenden zwei Bedingungen: zum einen werden die beiden Medikamente A und B in unterschiedlichem Umfang in den beiden Kliniken eingesetzt, z.B. wurde Medikament A an 80 Patienten in

Klinik I und an 120 Patienten in Klinik II eingesetzt. Zum anderen fällt in diesen Kliniken der Erfolg bzw. der Misserfolg der beiden Medikamente unterschiedlich aus. Die absolute Anzahl der Erfolge in Klinik I bei Medikament A beträgt 60, während diese Zahl bei Medikament B bei 100 liegt. In Klinik II nimmt die Anzahl der Erfolge bei Medikament A den Wert 40 an, wohingegen das Medikament B in der gleichen Klinik den Wert 10 annimmt. Durch Kombination der beiden Bedingungen entsteht bei der Aggregation der beiden Kliniken eine bessere Erfolgsquote bei Medikament B wohingegen bei der Einzelbetrachtung Medikament A die bessere Erfolgsquote hat:

$$\text{Klinik I: } P(\text{Erfolg}|A) = \frac{60}{80} = 0.75 \text{ und } P(\text{Erfolg}|B) = \frac{100}{150} = 0.66667$$

$$\text{Klinik II: } P(\text{Erfolg}|A) = \frac{40}{120} = 0.3333 \text{ und } P(\text{Erfolg}|B) = \frac{10}{40} = 0.25$$

Aggregierte Betrachtung der Kliniken I und II:

$$P(\text{Erfolg}|A) = \frac{60+40}{80+120} = \frac{100}{200} = 0.50 \text{ und}$$

$$P(\text{Erfolg}|B) = \frac{100+10}{150+40} = \frac{110}{190} = 0.5789.$$

**Umsetzung im Statistikkolabor:** Der Studierende bekommt die Möglichkeit, durch eine bereitgestellte Funktion die Zelleninhalte  $n_{ij}$  der beiden Kontingenztabelle von Klinik 1 und Klinik 2 beliebig zu verändern. Die aggregierte Kontingenztabelle wird von der Funktion automatisch berechnet und aufgestellt. Außerdem berechnet sie aus den bedingten relativen Häufigkeiten die dazugehörigen Erfolgsquoten der einzelnen Kliniken und der dritten Kontingenztabelle, in der die beiden Kliniken zusammen betrachtet werden. Durch diese Möglichkeit der Interaktion kann der Studierende nicht nur den Zustand betrachten, in dem das Simpson-Paradox auftritt, sondern zusätzlich beliebig von ihm selbst konstruierte Zustände, bei denen es nicht auftritt. Das konstruktivistische Prinzip der Wissenskonstruktion durch Aktivität wird angewendet. Der Studierende kann durch leichte Variation der  $n_{ij}$  erkennen, ab wann die Erfolgsquote der aggregierten Kontingenztabelle kippt, das Simpson-Paradox eintritt - "Learning-by-doing" findet statt. Die dem Studierenden bereitgestellte Funktion ist so voreingestellt, dass sie in den ersten drei Tabellen das Simpson-Paradox demonstriert und in den letzten drei Tabellen zeigt, wie dieses durch eine leichte Variation der absoluten Häufigkeiten  $n_{ij}$  der einzelnen Zellen nicht mehr auftritt. Betrachten wir die folgende Abbildung: die Funktion ist so voreingestellt, dass die beiden Werte 40 und 80<sup>111</sup> so variiert werden (40 wird zu 58 und 80 zu 62), dass sich das Simpson Paradox in den unteren drei Tabellen (von denen der Leser auf der Abbildung 13 nur die erste sehen kann) nicht mehr einstellt.

---

<sup>111</sup>Diese gehören zur zweiten Kontingenztabelle.

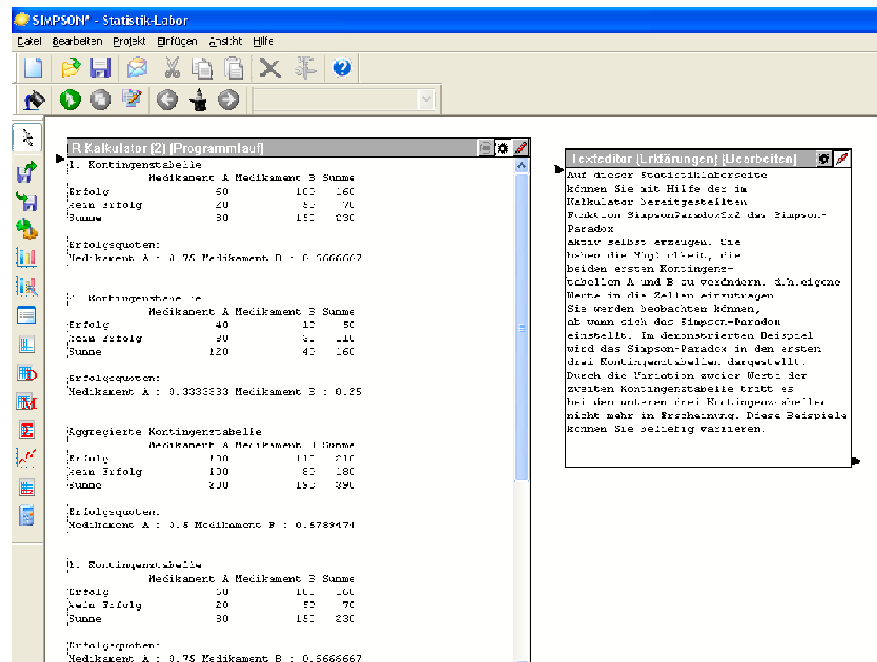


Abbildung 12: Simpson-Paradox: Interaktivität im Statistiklabor

**Technische Details zu den Statistiklaborseiten** Die Variation der  $n_{ij}$  findet in der bereitgestellten Funktion „SimpsonParadox2x2“ statt. Die Werte, die in der folgenden Grafik durch die Elipse gekennzeichnet sind, können vom Studierenden variiert werden. Es findet alles in einem R-Kalkulator statt. Dieser kann zwei Zustände annehmen. In der voranstehenden Grafik ist er dem Leser in seiner Run-Funktion begegnet, d.h. die Funktion wurde bereits ausgeführt bzw. der Programmablauf hat stattgefunden, die Kontingenztabelle waren zu sehen. In der folgenden Grafik befindet sich der R-Kalkulator in seiner Edit-Funktion, d.h. er kann bearbeitet werden. In diesem Zustand erfolgen die Programmierung sowie die Eingabe neuer Werte.

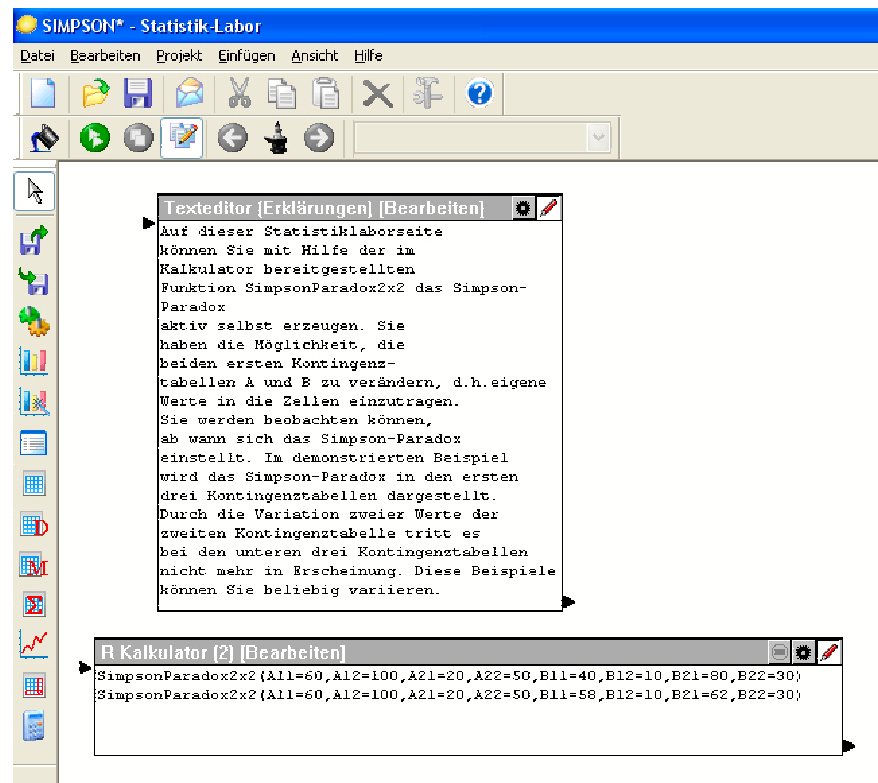


Abbildung 13: Variationsmöglichkeit im R-Kalkulator

Innerhalb des Funktionsaufrufs werden die Zelleninhalte der zwei Kontingenztafeln festgelegt. Die ersten vier Werte sind die Zelleninhalte  $n_{11}$  (60),  $n_{12}$  (100),  $n_{21}$  (20) und  $n_{22}$  (50) der ersten Kontingenztafel in entsprechender Reihenfolge. Entsprechendes gilt für die nächsten vier Werte der zweiten Kontingenztafel. Die Parameter, die hinter den Zelleninhalten erscheinen, bezeichnen die Ausprägungen der in der Kontingenztafel enthaltenen Merkmale.

Die demonstrierte Laborseite verdeutlicht die Möglichkeiten des Statistiklabors. Durch die Benutzung des R-Kalkulators kann der Studierende auf dessen Funktionalitäten aufmerksam gemacht werden.

**Bemerkung zum Einsatz** Im Evaluationsprozess gesammelte Erfahrungen haben gezeigt, dass eine reine Bereitstellung solcher Statistiklaborseiten dazu führen kann, dass statt einer Wissenskonstruktion im Sinne des Konstruktivismus eher ein "Herumspielen" seitens des Studierenden stattfindet, d.h. self-paced-learning hat nicht funktioniert. Aus diesem Grunde schlagen wir die Einbindung des Lernmoduls "Simpson-Paradox" in eine Statistikveranstaltung vor, d.h. die Komponenten des Lernmoduls werden in der Vorlesung, der Übung und vom Studierenden allein verwendet. Der Konstruktivismus nennt dieses Prinzip "guided learning". Das Ziel besteht darin, den Studierenden zur Wissenskonstruktion anzuregen. In der Vorlesung wird demonstriert und erklärt, in der Übung vertieft und erklärt, zu Hause vertieft und experimentiert der Studierende. Dies geschieht anhand verschiedener Beispiele, um die Interdisziplinarität der Statistik zu untermauern.

**Das Simpson-Paradox - wahrscheinlichkeitstheoretisch** Neben der experimentellen Erklärung und der Argumentation am Beispiel soll eine wahrscheinlichkeitstheoretische<sup>112</sup> erfolgen. Diese kann in Form einer Theoriekomponente in dem multimedialen Lernmodul *Simpson-Paradox* erscheinen. Das Simpson-Paradox wird als Paradox der bedingten Wahrscheinlichkeit erklärt. Dem Studierenden wird der Zusammenhang zwischen Kontingenztabelle und bedingten Häufigkeiten vermittelt. Er lernt außerdem den Umgang mit bedingten Häufigkeiten.

Es existieren drei Ereignisse A, B und C, für die wir die folgenden Annahmen treffen:

$$\begin{aligned} P(A|B) &< P(A|\bar{B}), \\ P(A|BC) &> P(A|\bar{B}C) \quad \text{und} \\ P(A|B\bar{C}) &< P(A|\bar{B}\bar{C}). \quad \text{Außerdem gilt:} \\ P(A|B\bar{C}) &< P(A|\bar{B}C). \end{aligned}$$

Um uns die Bedeutung dieser Annahmen vor Augen zu führen, folgt ein Beispiel. Die Ereignisse A, B und C sollen die folgenden Bedeutungen haben:

A: Es ist Stau.

B: Es ist Sonntag.

C: Es ist letzte Ferienwoche.

Es handelt sich nicht um ein empirisch bewiesenes Beispiel, sondern um ein fiktives, das der Verdeutlichung der Argumentation dient.

Betrachtet man  $P(A|B)$  als **Mittel** von  $P(A|BC)$  und  $P(A|B\bar{C})$ , so bedeutet dies: Die Wahrscheinlichkeit für Stau unter der Bedingung, dass Sonntag ist ( $P(A|B)$ ), kann mit Hilfe der folgenden bedingten Wahrscheinlichkeiten ermittelt werden:

- der Wahrscheinlichkeit für Stau unter der Bedingung, dass Sonntag und letztes Ferienwochenende ist ( $P(A|BC)$ ) und
- der Wahrscheinlichkeit für Stau unter der Bedingung, dass Sonntag und nicht letztes Ferienwochenende ist ( $P(A|B\bar{C})$ ).

Analog kann  $P(A|\bar{B})$  als Mittel von  $P(A|\bar{B}C)$  und  $P(A|\bar{B}\bar{C})$  betrachtet werden: Die Wahrscheinlichkeit für Stau unter der Bedingung, dass nicht Sonntag ist ( $P(A|\bar{B})$ ), kann als Mittel der folgenden bedingten Wahrscheinlichkeiten betrachtet werden:

- der Wahrscheinlichkeit für Stau unter der Bedingung, dass nicht Sonntag und letztes Ferienwochenende ist ( $P(A|\bar{B}C)$ ) und
- der Wahrscheinlichkeit für Stau unter der Bedingung, dass nicht Sonntag und nicht letztes Ferienwochenende ist ( $P(A|\bar{B}\bar{C})$ ).

<sup>112</sup>in Anlehnung an [Szekely1990], S. 63ff.

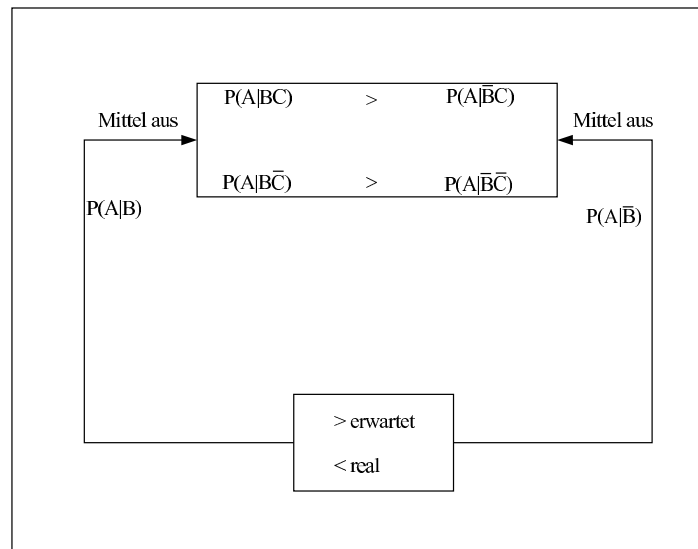


Abbildung 14: Wahrscheinlichkeitstheoretische Überlegungen zum Simpson-Paradox

Betrachtet man diese Zusammenhänge, dann tritt ein Paradox auf. Aufgrund der in der Grafik dargestellten Beziehungen muss das Mittel aus  $P(A|\bar{B}C)$  und  $P(A|\bar{B}\bar{C})$  kleiner sein als das Mittel aus  $P(A|BC)$  und  $P(A|B\bar{C})$ . Es müsste also gelten:

$$P(A|\bar{B}) < P(A|B)$$

Der Denkfehler besteht darin, dass der Leser das gewogene Mittel gleichsetzt mit dem arithmetischen Mittel. Bei der Bildung des gewogenen Mittels werden zur Bestimmung von  $P(A|B)$  und  $P(A|\bar{B})$  **verschiedene Gewichte** verwendet - je nach bedingter Wahrscheinlichkeit:

$$P(A|B) = \mathbf{P}(C|B)P(A|BC) + \mathbf{P}(\bar{C}|B)P(A|B\bar{C})$$

und

$$P(A|\bar{B}) = \mathbf{P}(C|\bar{B})P(A|\bar{B}C) + \mathbf{P}(\bar{C}|\bar{B})P(A|\bar{B}\bar{C})$$

Wir sehen, dass bei  $P(A|B)$  und  $P(A|\bar{B})$  die gewogenen Mittel nicht miteinander übereinstimmen.

Das Simpson-Paradox tritt allerdings nicht auf, wenn die Ereignisse B und C unabhängig sind, denn dann gilt:

$$P(C|B) = P(C|\bar{B})$$

und

$$P(\bar{C}|B) = P(\bar{C}|\bar{B})$$

Das Beispiel muss an dieser Stelle nicht weiterentwickelt werden, da es lediglich dazu diente, die Zusammenhänge der drei Ereignisse zu verdeutlichen, insbesondere die Argumentation über gewogene Mittel.

Um den Studierenden den Zusammenhang zwischen diesen wahrscheinlichkeitstheoretischen Ausführungen und der in den anderen Komponenten geführten Argumentation über Kontingenztabelle zu verdeutlichen, bietet es sich an, die Situation in Form zweier 2x2-Kontingenztabelle darzustellen:

	A Stau:	
	B Sonntag	$\bar{B}$ nicht Sonntag
C letzte Ferienwoche	Werte	beliebig
$\bar{C}$ nicht letzte Ferienwoche		

	$\bar{A}$ kein Stau:	
	B Sonntag	$\bar{B}$ nicht Sonntag
C letzte Ferienwoche	Werte	beliebig
$\bar{C}$ nicht letzte Ferienwoche		

**Das Simpson-Paradox - reale Beispiele** In diesem Abschnitt werden verschiedene Beispiele aus der traditionellen Statistikausbildung beschrieben, die das Auftreten des Simpson-Paradoxes im realen Leben behandeln.

- Ein historisches Beispiel zeigt das tatsächliche Auftreten des Simpson-Paradoxes anhand von Daten aus dem Jahr 1910 in den U.S.A.
- Ein Beispiel aus dem Sport beschreibt das Auftreten des Simpson-Paradoxes im Basketball und könnte zur Motivation der Studierenden eingesetzt werden.

Jedes dieser Beispiele könnte in einer Beispielkomponente der Lernsoftware *Statistik interaktiv komplett* umgesetzt werden und einen Teil des Lernmoduls *Simpson-Paradox* darstellen. Es kommt also zu einer Kombination traditioneller und multimedialer Vermittlungsmethoden, d.h. die Idee der Vermittlung des Simpson-Paradoxes anhand realer Daten wird in dem Lösungsvorschlag aufgenommen.

Eine weitere Möglichkeit des Umgangs mit den Beispielen besteht darin, gemeinsam mit den Studierenden eines dieser Beispiele im Statistikkolabor analog zu der in der Vorlesung demonstrierten Statistikkolaborseite umzusetzen. Die Studierenden würden schrittweise die Arbeit des R-Kalkulators verstehen, der Tutor könnte einzelne Befehle genauer erklären. Zudem würde das Zustandekommen des Simpson-Paradoxes wiederholt verdeutlicht werden. In der Übung könnten dann Werte der Kontingenztafeln variiert werden und man könnte gemeinsam diskutieren, wieso und ab wann das Paradox auftritt.

**Historisches Beispiel:** Im Jahr 1910 hatte die Welt mit einer Tuberkulose-Epidemie zu kämpfen. In diesem Zusammenhang wurden die Tuberkulose-Toten der Städte New York und Richmond dokumentiert. Dabei wurde bezüglich der Hautfarbe (*weiß oder farbig*) differenziert. Die Ergebnisse sind in den beiden folgenden Kontingenztafeln dargestellt:

weiß	Tuberkulose-tote	$\overline{\text{Tuberkulose-tote}}$	Summe
	New York	8365	4666809
Richmond	131	80764	80895
Summe	8496	4747573	4756069

farbig	Tuberkulose-tote	$\overline{\text{Tuberkulose-tote}}$	Summe
	New York	513	91196
Richmond	155	46578	46733
Summe	668	13774	14442

weiß und farbig	Tuberkulose-tote	<i>Tuberkulose-tote</i>	Summe
New York	8878	4758005	4766883
Richmond	286	127342	127628
Summe	9164	4885347	4894511

Vergleicht man die Sterblichkeitsraten (bezogen auf Tuberkulose) beider Städte, so ist die New Yorks (0.00186) geringer als die Richmonds (0.00224). Bildet man allerdings zwei Kategorien, d.h. unterteilt man die Gesamtbevölkerung beider Städte in Weiße und Farbige, so kommt man zu einem anderen Ergebnis. Sowohl für Weiße als auch für Farbige liegt einzeln betrachtet die Sterblichkeitsrate Richmonds (für Weiße: 0.00162, für Schwarze: 0.00332) unter der New Yorks (für Weiße: 0.00179, für Schwarze: 0.00560). Das Simpson-Paradox tritt auf:

”If you are white, go to Richmond. If you are colored, go likewise to Richmond.  
If you are white or colored, then remain in New York.  
You don’t have to falsify statistics—they *are* falsifications.”<sup>113</sup>

Analog zum Beispiel aus der Medizin kann das Simpson-Paradox in dieser Situation erklärt werden:

Im Wesentlichen liegt die Begründung für das Paradox im Auftreten zweier Bedingungen. Zum einen lebten in New York und Richmond unterschiedlich viele Personen der beiden Hautfarben. Zum anderen ist bei separater Betrachtung der Hautfarben die Anzahl der Todesfälle an Tuberkulose in den beiden Städten unterschiedlich.

Motiviert der Kontext dieser Daten die Studierenden nicht, kann das folgende Beispiel eingesetzt werden.

**Basketball:** Beginnen wir die Argumentation mit einem Zitat, das die These untermauert, sich bei der Verbesserung der Statistikausbildung der Vermittlung statistischer Paradoxe zu widmen:

”This paper gives two recommendations to researchers and teachers. First, take care to consider what data are available to laypersons. Second, it is important to make the public aware of Simpson’s paradox and other counterintuitive results.”<sup>114</sup>

Basketballfans sprechen von dem ”Hot Hand” Phänomen welches besagt, dass ein Spieler bei einem Freiwurf<sup>115</sup> nach einem Treffer besser wirft als nach einem Wurf ins Leere. Wardrop greift Arbeiten von Kahneman, Slovic und Tversky auf und untersucht, ob die Entscheidungen, die der Fan trifft, mit denen eines Statistikers übereinstimmen. Die Kernfrage lautet:

”When shooting free throws, does a player have a better chance of making his second shot after making this first shot than after missing his first shot?”<sup>116</sup>

<sup>113</sup>[SimpsonRichmond].

<sup>114</sup>[Wardrop1995].

<sup>115</sup>Ein Freiwurf besteht aus zwei aufeinander folgenden Würfen.

<sup>116</sup>[Wardrop1995], S. 24.



68 Prozent der Fans beantworten diese Frage positiv, stimmen also für das "Hot-Hand" Phänomen. Anhand zweier Spieler der "Boston Celtics" demonstriert Wardrop das Simpson-Paradox, indem er dieser Frage nachgeht. Er stellt eine Kontingenztabelle für Larry Bird, einen erfolgreichen Spieler der "Boston Celtics" und eine für Rick Robey, einen erfolglosen Spieler auf:

Larry Bird	Treffer2	Miss2	Summe
Treffer1	251	34	285
Miss1	48	5	53
Summe	299	39	338

Rick Robey	Treffer2	Miss2	Summe
Treffer1	54	37	91
Miss1	49	31	80
Summe	103	68	171

Die Kontingenztabelle werden beispielhaft für Bird interpretiert: Larry Bird hat 338-mal einen Freiwurf geworfen. 251-mal hat er beide getroffen, 5-mal keinen. 34-mal war nur der erste Wurf ein Treffer, 48-mal nur der zweite. Um das "Hot-Hand" Phänomen zu untersuchen, vergleicht man die Wahrscheinlichkeiten dafür, dass der erste Wurf ein Treffer war und gefolgt wurde von einem weiteren Treffer  $p_{hit}$  oder dass der erste Wurf kein Treffer (Miss) war und gefolgt wurde von einem Treffer  $p_{miss}$ . Dies erfolgt über die Berechnung der bedingten relativen Häufigkeiten:

$$\text{Larry Bird: } p_{hit} = \frac{251}{285} = 0.881 \text{ und } p_{miss} = \frac{48}{53} = 0.906$$

$$\text{Rick Robey: } p_{hit} = \frac{54}{91} = 0.593 \text{ und } p_{miss} = \frac{49}{80} = 0.612$$

Diese statistischen Ergebnisse sprechen eindeutig gegen das "Hot-Hand" Phänomen der Basketballfans, auch wenn diese Ergebnisse nicht als statistisch signifikant bezeichnet werden können. Aber jeder der beiden Spieler traf nach einem "Miss" besser als nach einem Treffer. Das "Hot-Hand" Phänomen besagt genau das Gegenteil.

Betrachten wir nun die aggregierte Kontingenztabelle dieser beiden Spieler, so tritt das Simpson-Paradox auf:

Beide Spieler	Treffer2	Miss2	Summe
Treffer1	305	71	376
Miss1	97	36	133
Summe	402	107	509

Dies wird an der Berechnung der bedingten relativen Häufigkeiten deutlich:

$$p_{hit} = \frac{305}{376} = 0.811 \text{ und } p_{miss} = \frac{97}{133} = 0.729$$

In der aggregierten Kontingenztabelle gilt:  $p_{hit} > p_{miss}$ , während in den Kontingenztabelle der einzelnen Spieler genau das Gegenteil gilt:  $p_{hit} < p_{miss}$ .

Die Begründung für das Paradox liegt im Auftreten der folgenden zwei Bedingungen. Zum einen erzielten Larry Bird und Rick Robey beim ersten Wurf unterschiedlich viele Treffer. Zum anderen ist bei separater Betrachtung der Spieler die Anzahl der Treffer im ersten und zweiten Wurf unterschiedlich. Bei der gemeinsamen Betrachtung beider Spieler tritt dadurch das "Hot-Hand" Phänomen auf.

Das Beispiel eignet sich neben der Vermittlung des Simpson-Paradoxes dazu, den Studierenden zu demonstrieren, wo überall statistische Kenntnisse weiterhelfen, selbst in Situationen, in denen sie es nie erwartet hätten.

**Fiktives, wirtschaftswissenschaftliches Beispiel** Ein selbstkonstruiertes wirtschaftswissenschaftliches Beispiel zeigt ein mögliches Auftreten des Simpson-Paradoxes in Unternehmen und stellt den Bezug der Statistik zum späteren Berufsleben der Studierenden dar. Anhand des Beispiels wird gezeigt, dass das Simpson-Paradox nicht nur in 2x2-Kontingenztabelle auftreten kann, sondern auch in höheren Dimensionen.

Ein Unternehmen hat zwei Produktionsstätten, 1 und 2. In beiden produziert es Schrauben auf drei verschiedenen Maschinentypen, A, B und C. Die Unternehmensleitung beschließt, die Produktion des Unternehmens auszudehnen. Zu diesem Zweck soll für beide Produktionsstätten jeweils eine neue Maschine angeschafft werden. Die Unternehmensleitung steht nun vor der Entscheidung, welcher der drei Maschinentypen neu angeschafft werden soll. Als Entscheidungskriterium sollen die Ausschussmengen der einzelnen Maschinen herangezogen werden. Es wird der Maschinentyp neu angeschafft, der die geringste Ausschussmenge produziert. In dieser Situation kann das Simpson-Paradox auftreten: ist zum Beispiel in beiden Produktionsstätten die Maschine des Typs A zuverlässiger bezüglich der gemessenen Ausschussquote, läge die Entscheidung nahe, zwei neue Maschinen dieses Typs A anzuschaffen. Das Simpson-Paradox tritt in diesem Zusammenhang auf, wenn die aggregierte Betrachtung aller drei Maschinentypen zu dem Ergebnis kommt, dass der Maschinentyp A nicht die geringste Ausschussmenge produziert. Betrachten wir hierzu zunächst die Kontingenztabelle der beiden Produktionsstätten einzeln und anschließend zusammen, d.h. die aggregierte Kontingenztabelle:

<b>Produktionsstätte 1</b>	Ausschuss	kein Ausschuss	Summe
Maschine A	20	6000	6020
Maschine B	50	10000	10050
Maschine C	20	600	620
Summe	90	16600	16690

<b>Produktionsstätte 2</b>	Ausschuss	kein Ausschuss	Summe
Maschine A	80	4000	4080
Maschine B	30	1000	1030
Maschine C	33	488	521
Summe	143	5488	5631

Beide Produktionsstätten	Ausschuss	kein Ausschuss	Summe
Maschine A	100	10000	10100
Maschine B	80	11000	11080
Maschine C	53	1088	1141
Summe	233	22321	22321

Anhand der 3x2-Kontingenztafeln ist zu erkennen, dass in den beiden Produktionsstätten unterschiedliche Ausschussmengen realisiert werden. Außerdem unterscheiden sich die auf den verschiedenen Maschinentypen produzierten Ausschussmengen. Betrachtet man die Ausschussquote der verschiedenen Maschinen innerhalb der jeweiligen Produktionsstätte, so ist das Simpson-Paradox aufgetreten. Die folgende Statistikkalorseite berechnet die Ausschussquoten der einzelnen Maschinen über bedingte relative Häufigkeiten.

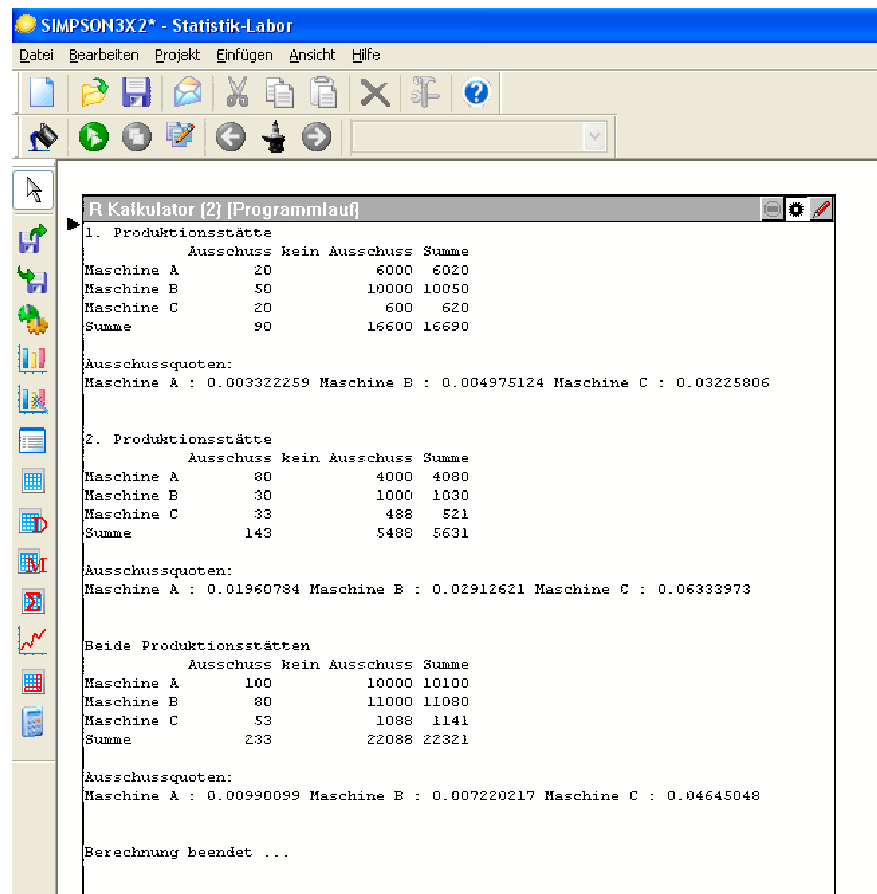


Abbildung 15: Simpson-Paradox in einer 3x2-Kontingenztafel

Dem Studierenden wird die Möglichkeit der Interaktivität gegeben, um das Zustandekommen des Simpson-Paradoxes experimentell nachzuvollziehen:<sup>117</sup>.

<sup>117</sup>Dies geschieht analog zu dem oben demonstrierten Beispiel der 2x2-Kontingenztafel.

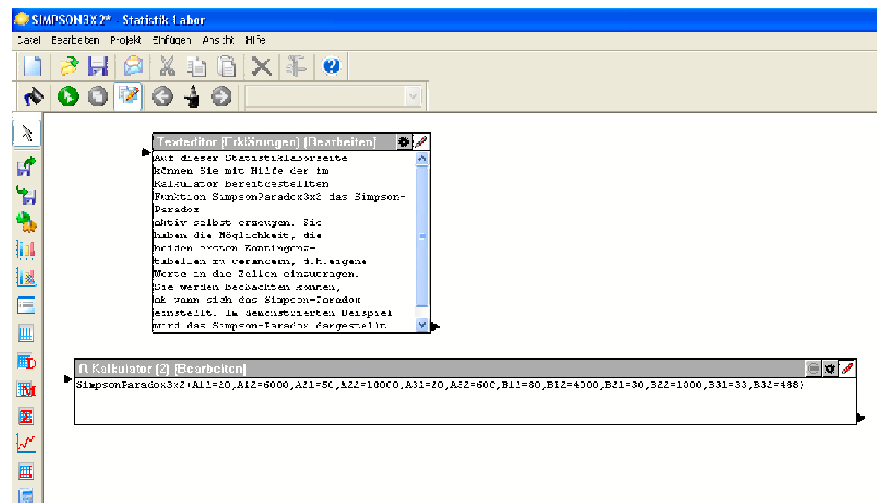


Abbildung 16: Variationsmöglichkeit im R-Kalkulator

Die Erklärung des Simpson-Paradoxes erfolgt analog zu den anderen Beispielen. Der Studierende kann anhand dieses Beispiels das Simpson-Paradox erklären: es kann Wissenskonstruktion im Sinne des Konstruktivismus stattfinden, indem das vorhandene Wissen über das Paradox auf dieses Beispiel angewendet wird.

**Bezug zur Defizitklassifikation** Im Folgenden werden die Defizite <sup>118</sup>, denen in diesem Lösungsvorschlag entgegengewirkt werden soll, näher betrachtet:

- Defizit B (Interaktivität): Die entwickelte Statistiklaborseite zum Simpson-Paradox ermöglicht und fördert die Interaktivität des Studierenden - er ist aktiv Lernender, eignet sich das Verständnis bezüglich des Simpson-Paradoxes durch eigenes Experimentieren an.
- Defizit C (Interdisziplinarität und Realitätsbezug): Der Studierende wird durch die gewählten Beispiele aus den Bereichen Medizin, Sport und Wirtschaftswissenschaften auf den interdisziplinären Charakter der Statistik aufmerksam. Durch die Vermittlung des Simpson-Paradoxes anhand realer Daten wird der Bezug zwischen der Statistik und der Realität bzw. dem späteren Berufsleben des Studierenden hervorgehoben. Es werden Brücken zwischen verschiedenen Disziplinen zwischen der Realität und der Statistik geschlagen. Man könnte sich fragen, wo die Besonderheit der neuen Medien liegt, da die Beispiele aus der traditionellen Statistikausbildung stammen. Der Vorteil der neuen Medien könnte darin bestehen, dass in einem Lernmodul "Simpson's Paradox" jedes der später vorgestellten Beispiele in Form einer Beispielkomponente eingebracht werden kann. Wie in der traditionellen Statistikausbildung könnte anhand eines Beispiels in der Vorlesung argumentiert werden. Die Motivation der Studierenden könnte durch den Verweis auf die weiteren Beispiele aus den unterschiedlichsten Disziplinen geweckt werden. Auf diese Weise könnten die Studierenden durch eigene Aktivität, d.h. eigenes Bearbeiten der weiteren Beispiele (dies könnte auch durch eine Übung unterstützt werden) Interdisziplinarität

<sup>118</sup>Vgl. 3.

und Realitätsbezug der Statistik selbst erfahren. Die neuen Medien bieten also die Möglichkeit, viele reale Beispiele unterschiedlicher Disziplinen als eine Einheit in einem multimedialen Lernmodul "Simpson's Paradox" zu präsentieren. Exemplarisch werden in diesem Kapitel drei Beispiele demonstriert, die jeweils eine Beispielkomponente bilden könnten. An dieser Stelle wird deutlich, dass Ansätze der traditionellen Statistikausbildung aufgegriffen und erweitert werden, nicht zwingend das Rad neu erfunden wird.

- Defizit E (Paradoxe): Um Verständnis für das Simpson-Paradox zu bekommen, wird sein Zustandekommen erklärt und experimentell erlernbar gemacht. Die genaue Erklärung über die bedingten relativen Häufigkeiten und der Verhältnisse dieser zueinander konzentriert sich auf die Vermittlung möglicher Ursachen des Paradoxes, geht somit weit über dessen bloße Darstellung hinaus. Zusammen mit der Interaktivität im Statistiklabor kann der Studierende erkennen, wie das Simpson-Paradox zustandekommt.
- Defizit G (Computational Statistics): die Vorteile von "Computational Statistics", besonders die zusätzlichen Möglichkeiten, die durch den Einsatz des Rechners bestehen, werden durch den Lösungsvorschlag hervorgehoben. Erst der Rechner, hier die Arbeit im Statistiklabor, ermöglicht die Interaktivität des Studierenden, fördert sein Verständnis des Paradoxes. Des Weiteren wird der Studierende durch das Experimentieren im Statistiklabor an dieses herangeführt. In einem zweiten Schritt könnte er selbst Statistiklaborseiten erstellen und die Rolle des Rechners in der Statistik selbst erkennen.
- Die wahrscheinlichkeitstheoretischen Ausführungen sollen eine Verbindung zwischen Kontingenztabelle und bedingten Wahrscheinlichkeiten und dem mathematischen Umgang mit diesen aufzeigen - es wird dem Defizit H (fehlendes Verständnis für Zusammenhänge statistischer Methoden) entgegengewirkt.

Der Lösungsvorschlag zeigt, wie diese Defizite behoben werden können, wie die Statistikausbildung bezüglich dieser Defizite positiv beeinflusst werden kann. Methoden der traditionellen Statistikausbildung, wie z.B. der Einsatz realer Daten, werden kombiniert mit den Möglichkeiten der neuen Medien.

**Zwischenfazit:** Der präsentierte Lösungsvorschlag kombiniert Möglichkeiten der traditionellen und multimedialen Statistikausbildung. Die Idee der traditionellen Statistikausbildung bezüglich der Verwendung realer Daten und Beispiele wird erweitert durch die Möglichkeit der Interaktivität im Statistiklabor. Der Studierende lernt innerhalb eines in sich geschlossenen Lernmoduls *Simpson-Paradox*. Dieses setzt sich aus den Komponenten Theorie, Statistiklabor, Beispiele zusammen. Im Statistiklabor besteht die Möglichkeit der Interaktivität des Studierenden. Er kann das Simpson-Paradox experimentell sowohl bei einer 2x2 als auch bei einer 2x3 Kontingenztabelle herbeiführen.<sup>119</sup>

---

<sup>119</sup>Die Beispiele wurden nicht in *Statistik interaktiv komplett* umgesetzt, sondern im Dokument beschrieben, da der Leser durch diese Darstellung mehr sieht als auf Screenshots möglicher Beispielkomponenten.

### 4.3 Strukturen in Daten

Den Ausgangspunkt des hier dargestellten Lösungsvorschlags bildet ein empirischer Datensatz. Das Thema ist die Suche nach Strukturen in den Daten. Dabei können z.B. folgende Probleme auftreten:

- Ohne den Einsatz des Rechners ist der Arbeitsaufwand sehr hoch, da die Anwendung vieler Verfahren notwendig ist.
- Pro Verfahren macht sich zusätzlich der Umfang des Datensatzes bemerkbar.

Die Struktur der hier vorliegenden Daten weist auf einen nicht normalverteilten Charakter hin. Das Ziel besteht darin, den Daten durch eine geeignete Transformation eine annähernd normalverteilte Gestalt zu geben. Zunächst sollen die Studierenden lernen, Strukturen in Daten zu erkennen. Dann sollen sie verstehen, dass die Struktur durch die Transformation der Daten umgewandelt werden kann. Außerdem soll ihnen vermittelt werden, dass eine Transformation von Daten neue Möglichkeiten der statistischen Datenanalyse mit sich bringen kann. Im Anschluss wird die Standardisierung beim statistischen Testen als Transformationsprozess beschrieben.

Durch Einsatz des Statistiklabors wird demonstriert, welche Möglichkeiten und Vorteile gegenüber der traditionellen Statistikausbildung bestehen.

**Daten, Ausgangspunkt, Defizit** Der Datensatz besteht aus dem Gehirngewicht von 62 Tieren.<sup>120</sup> Diese beinhalten den Menschen als Element der Klasse der Säugetiere. Es soll untersucht werden, ob der Mensch im Vergleich zu den anderen Tieren außergewöhnlich ist. Die Daten sind real und weisen der Größe nach sortiert einen exponentiellen Verlauf auf. Die Verteilung der Daten wird anhand des Histogramms und Boxplots auf der folgenden Statistiklaborseite dargestellt:

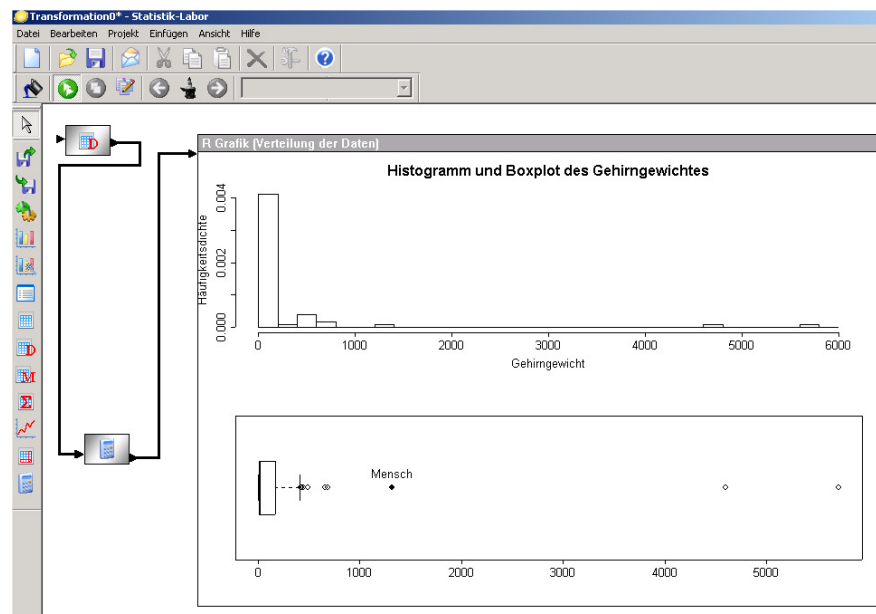


Abbildung 17: Histogramm des Gehirngewichtes

<sup>120</sup>Quelle: [www.fsmq.org/filelibrary/xls/auaasmammasi.xls](http://www.fsmq.org/filelibrary/xls/auaasmammasi.xls)

Ein Großteil statistischer Methoden verlangt Symmetrie der Daten. Die Normalverteilung ist eine symmetrische Verteilung, die häufig vorausgesetzt wird, um statistische Methoden anwenden zu können. Sowohl das Histogramm als auch der Boxplot zeigen, dass die Daten nicht symmetrisch, sondern rechtsschief verteilt sind. Die Merkmalsausprägung des Menschen ist im Boxplot gekennzeichnet. Aufgrund ihrer Lage könnte man den Menschen hinsichtlich seines Gehirngewichts als außergewöhnlich bezeichnen.

Der Studierende wird in der traditionellen Statistikausbildung mit der Tatsache konfrontiert, dass die Annahme normalverteilter Daten gelten muss, um bestimmte statistische Methoden anzuwenden. Das Ziel dieses Lösungsvorschlags besteht darin, einen Schritt früher anzusetzen, d.h. beim Problem, das es gilt, mit statistischen Methoden zu lösen.

Bei der Betrachtung der Grafiken stellt sich die Frage, ob der Mensch in Bezug auf das Gewicht seines Gehirns außergewöhnlich ist. Um diese zu beantworten, könnte z.B. ein statistischer Test durchgeführt werden. Dieser setzt allerdings die Normalverteilung voraus. Das Ziel besteht darin, den Studierenden zu vermitteln, dass die Struktur der Daten durch eine Transformation umgewandelt werden kann und dass dies statistisch zulässig ist. Es gilt außerdem, ihnen zu zeigen, dass in verschiedenen Situationen unterschiedliche Arten von Transformationen zu der Annahme der Normalverteilung bezüglich der transformierten Daten führen können. Weiterhin wird vermittelt, dass unter der Annahme der Normalverteilung statistische Tests unter Zuhilfenahme der Standardisierung durchgeführt werden können. Es wird in drei Schritten vorgegangen:

1. Transformation: Überführung der asymmetrischen Gestalt der Daten in eine symmetrische, insbesondere in eine annähernd normalverteilte Gestalt.
2. Standardisierung: Überführung der annähernd normalverteilten Gestalt der Daten in eine standardnormalverteilte.
3. Transformation und Standardisierung: Überblick über das gesamte Transformationsverfahren durch Zusammenführung der beiden oberen Schritte.

Die Umsetzung dieser drei Schritte findet im Statistiklabor statt.

**Zur Umsetzung im Statistiklabor** Auf einer Statistiklaborseite werden vier verschiedene Transformationsmöglichkeiten angeboten, die interaktiv auf den Datensatz angewendet werden können. Ein alternativer Datensatz wird bereitgestellt. Auch auf diesen können die vier Transformationsmöglichkeiten angewendet werden. Der Vergleich beider Datensätze verdeutlicht, dass es von dem Charakter der Daten abhängt, welche Transformation am besten geeignet ist, Daten in eine annähernd normalverteilte Gestalt zu überführen. In der traditionellen Statistikausbildung kann der Transformationsprozess realer Daten interaktiv nicht durchgeführt werden. In der Veranstaltung ist es dem Dozenten nicht möglich, Transformationen in Echtzeit zu demonstrieren, da der damit verbundene Rechen- und Zeitaufwand viel zu groß ist.

In der rechnergestützten Lehre besteht die Möglichkeit der grafischen Darstellung sowie der interaktiven Transformation der Daten während der Veranstaltung. Der Vorteil der neuen Medien wird in den folgenden Ausführungen deutlich: die neuen Medien können so entwickelt und gestaltet werden, dass sie zusätzlich zu den genannten Vermittlungsmöglichkeiten des Dozenten Interaktivität der Studierenden ermöglichen und hervorrufen. Die Studierenden lernen mit Hilfe der neuen Medien (hier im Statistiklabor), dass mit Hilfe geeigneter Transformationen die Struktur von Daten umgewandelt werden kann. Die

Transformation *Logarithmierung* führt in dem hier gewählten Beispiel dazu, dass die zunächst rechtsschief verteilten Daten in eine symmetrische Gestalt gebracht werden. Die Gestaltung der entsprechenden Laborseiten erfolgte wie auch in den voranstehenden Beispielen den konstruktivistischen Prinzipien der Wissenskonstruktion gemäß durch eigenes Experimentieren bzw. durch Interaktionsmöglichkeiten der Studierenden.

Ausgehend von der Frage, ob der Mensch in Bezug auf das Gewicht seines Gehirns außergewöhnlich ist, werden die transformierten Daten, die durch die Logarithmierung eine annähernd normalverteilte Gestalt besitzen, in dem zweiten Schritt standardisiert. Das Ziel besteht darin, den Studierenden zu vermitteln, dass zum Testen eine Transformation notwendig sein kann, die von der Normal- zur Standardnormalverteilung. Dieser Transformationsprozess wird in der traditionellen Statistikausbildung lediglich durchgeführt. Der multimediale Lösungsvorschlag fokussiert den eigentlichen Transformationsprozess und bildet diesen grafisch ab. Auf einer für diese Zwecke entwickelten Statistiklaborseite wird die Möglichkeit der Interaktivität gegeben, der Standardisierungsprozess der Daten kann vom Studierenden durchgeführt werden.

Abschließend werden auf einer Statistiklaborseite die beiden Schritte gemeinsam dargestellt. Dies hat zum Ziel, den gesamten Prozess als Einheit darzustellen. Die oben geschaffenen Interaktionsmöglichkeiten bestehen weiterhin. Der Studierende kann die verschiedenen Stufen des Prozesses experimentell nachvollziehen.

**Statistiklaborseiten - Demonstration und Interaktivität** Technische Vorbemerkung: Der multimediale Lösungsvorschlag setzt den Fokus auf das Statistiklabor. Es werden zwei Typen von Statistiklaborseiten entwickelt:

- **Demonstrationsseiten:** Auf diesen Seiten werden Möglichkeiten demonstriert, Strukturen in Daten zu erkennen.
- **Interaktivitätsseiten:** Hier steht die Interaktivität des Studierenden im Mittelpunkt. Er soll Transformationsprozesse durch eigenes Experimentieren im Statistiklabor vermittelt bekommen. Er wird im Sinne des Konstruktivismus als aktiv Lernender angesprochen und gefordert.

Die im Folgenden vorgestellten Statistiklaborseiten bauen aufeinander auf und können somit schrittweise in einer Statistikveranstaltung eingesetzt werden. Die auf den einzelnen Statistiklaborseiten eingesetzten Funktionen können sowohl mit den bereitgestellten als auch mit beliebigen Datensätzen aufgerufen werden.

Ausgehend von der Frage, ob die Annahme der Normalverteilung als Voraussetzung für die Anwendung eines statistischen Tests erfüllt ist, wird die Verteilung der Daten näher untersucht. Die folgende Statistiklaborseite enthält drei Grafiken: den QQNorm-Plot des Gehirngewichtes der Tiere, die zugehörige Dichtefunktion sowie den Boxplot.



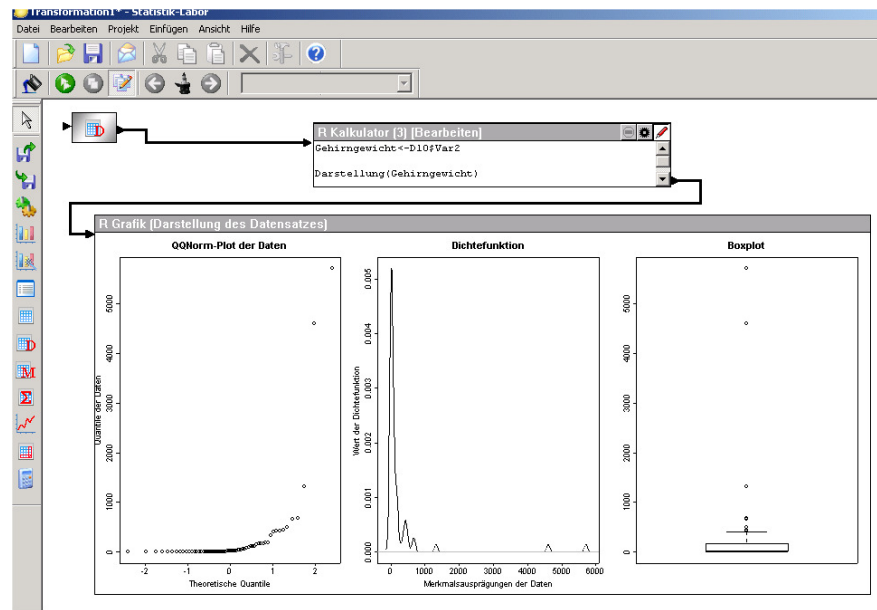


Abbildung 18: QQNorm-Plot, Dichtefunktion und Boxplot

Anhand des QQNorm-Plots wird deutlich, dass die Daten keine normalverteilte, sondern eine exponentielle Struktur aufweisen. Würden normalverteilte Daten vorliegen, würden die Punkte des QQNorm-Plots eine Gerade bilden. Auch die Dichtefunktion und der Boxplot sprechen gegen eine symmetrische Struktur der Daten. Die Gegenüberstellung der drei Methoden demonstriert verschiedene Möglichkeiten, auf die Daten zu blicken, um Aussagen über mögliche Strukturen in diesen treffen zu können. Die Studierenden werden auf die Notwendigkeit einer Transformation aufmerksam gemacht. Diese wandelt die Daten in eine symmetrische Struktur um.

Es stellt sich die Frage, welche Transformation für diese Zwecke am besten geeignet ist. Auf der folgenden Statistiklaborseite hat der Studierende die Möglichkeit zur Interaktivität. Er kann zwischen verschiedenen Transformationen wählen: Logarithmierung, Wurzelziehung, Quadrierung und Potenzierung dritten Grades. Es handelt sich um eine Auswahl gängiger Transformationen. Die Logarithmierung wurde ausgewählt, weil sie für die Umwandlung des ersten Datensatzes in eine symmetrische Gestalt gut geeignet ist. Gleiches gilt für die Wurzelziehung in Bezug auf den alternativ bereitgestellten Datensatz.

Das Ziel dieser Statistiklaborseite besteht darin, dem Studierenden verschiedene Transformationen und deren Eignung in Bezug auf die Symmetrisierung der Daten experimentell bzw. interaktiv zu vermitteln.

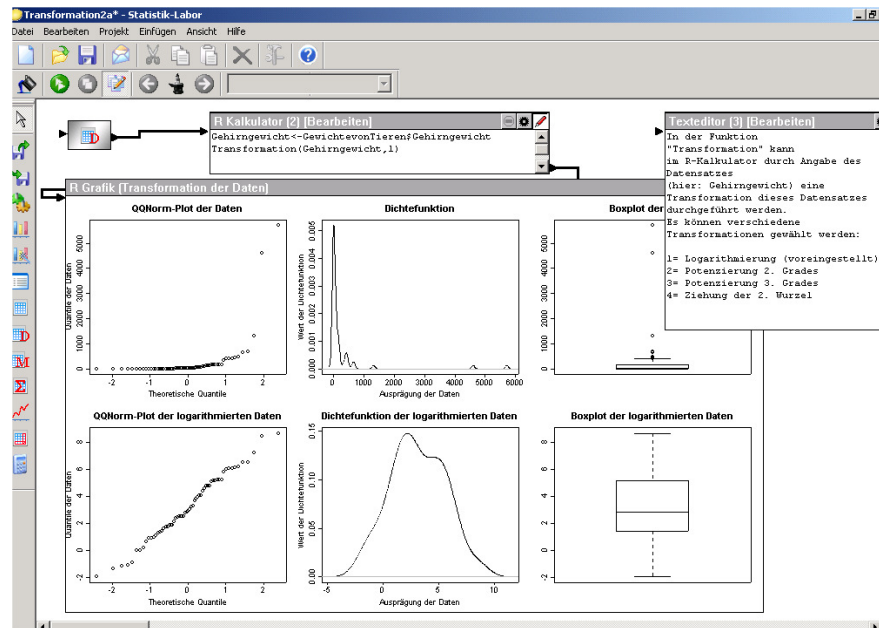


Abbildung 19: Interaktivität: Durchführung verschiedener Transformationen

Auf der Statistiklaborseite ist ein Texteditor enthalten, in dem der Umgang mit der Laborseite erklärt wird. Das R-Grafikfenster enthält sechs Grafiken, in der oberen Hälfte sind der QQNorm-Plot, die Dichtefunktion und der Boxplot der Daten dargestellt. In der unteren Hälfte der Grafik befinden sich diese Grafiken in Bezug auf die logarithmierten Daten. Durch diese Gegenüberstellung erkennt der Studierende die Wirkung der Transformation (hier: Logarithmierung):

- Der Vergleich der QQNorm-Plots zeigt, dass die Daten vor der Transformation einen exponentiellen, nach der Transformation einen linearen Verlauf aufweisen.
- Der Vergleich der Dichtefunktionen zeigt, dass auch der Verlauf der Dichtefunktion der Daten durch die Transformation deutlich verändert wird. Die Gestalt der Dichtefunktion nimmt nach der Transformation große Ähnlichkeit mit der der Normalverteilung an. Sie wird symmetrischer.
- Der Vergleich der Boxplots unterstützt diese Beobachtungen. Anhand des Boxplots der logarithmierten Daten ist ihre symmetrische Gestalt deutlich erkennbar.

Die folgende Statistiklaborseite zeigt, wie die Transformation für die Wurzelziehung 2. Grades aussieht:

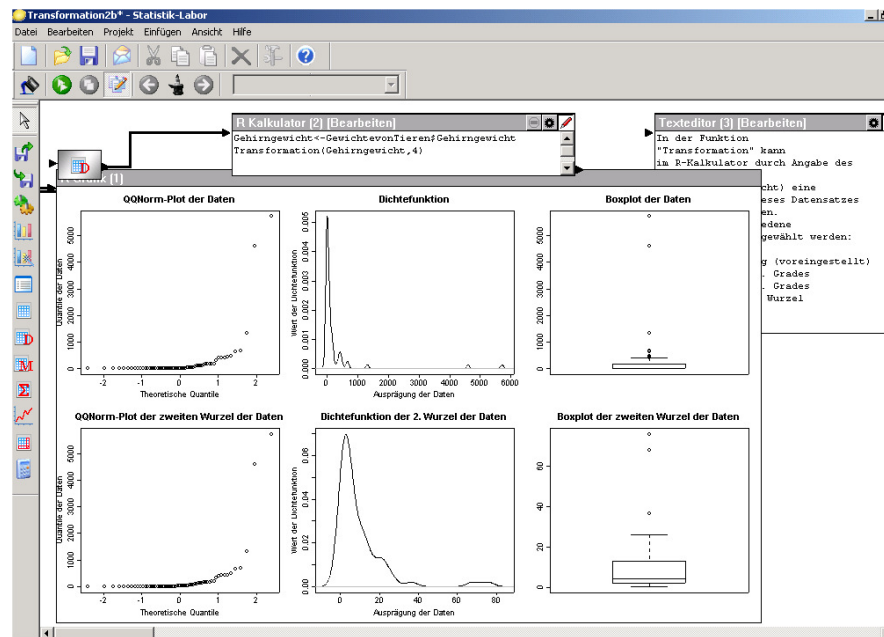


Abbildung 20: Interaktivität: Wurzelziehung 2. Grades

Die auf der folgenden Statistiklaborseite vorgenommene Gegenüberstellung mit der Logarithmierung zeigt, dass diese besser geeignet ist, die Daten in eine symmetrische Gestalt umzuwandeln:

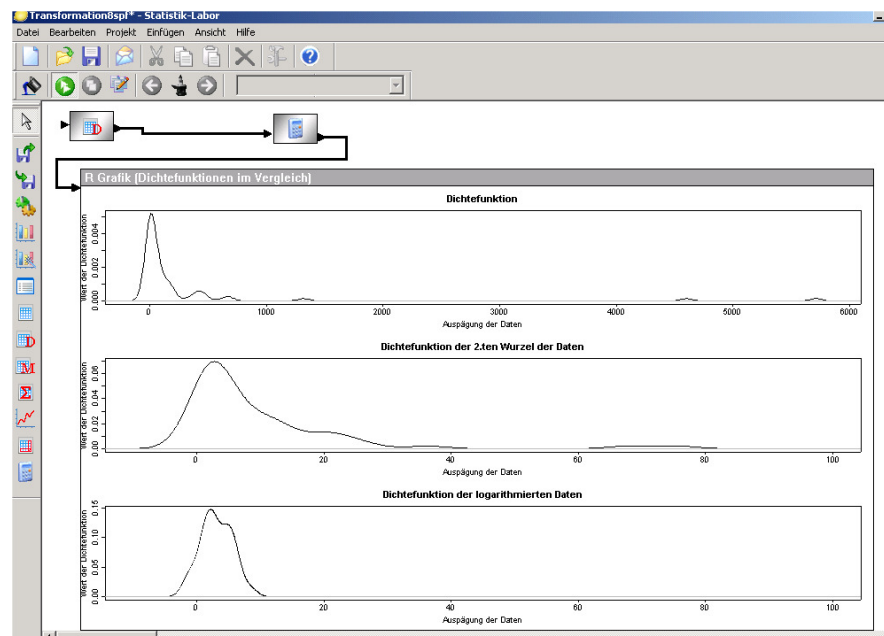


Abbildung 21: Gegenüberstellung zweier Transformationen

Bei dem Einsatz dieser Laborseite muss darauf hingewiesen werden, dass bei einem anderen Datensatz eine andere Transformation besser geeignet sein kann, die Daten in eine symmetrische Gestalt umzuwandeln.

**Standardisierung der Daten** Die folgende Statistiklaborseite geht einen Schritt weiter. Sie geht von den bereits transformierten (hier: logarithmierten) Daten aus und demonstriert eine weitere Transformation: die Standardisierung der Daten, wie sie beim Testen häufig benutzt wird.

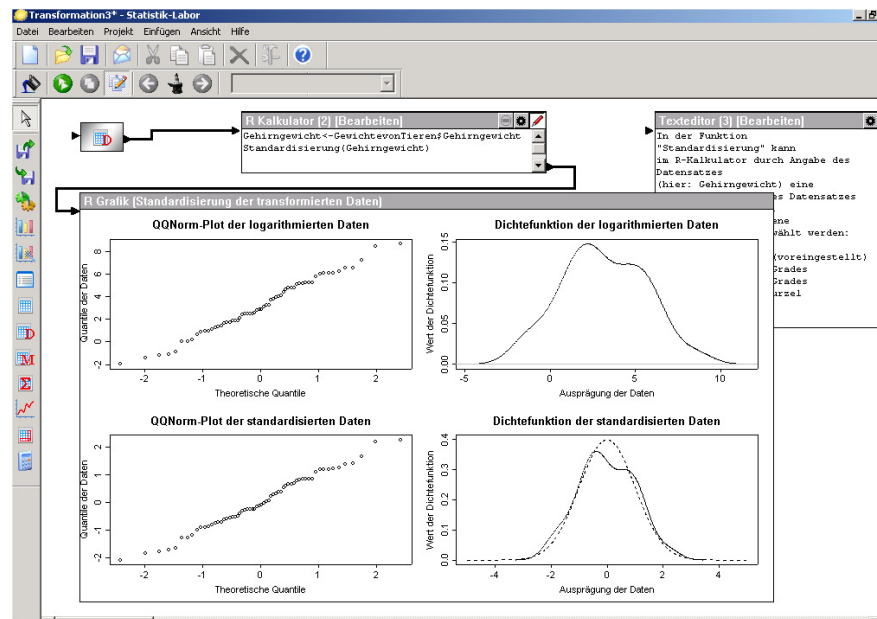


Abbildung 22: Dichtefunktionen im Vergleich

Im oberen Teil der Grafik sind der QQNorm-Plot und die Dichtefunktion der logarithmierten Daten dargestellt. Die Auswirkungen der Standardisierung auf die Gestalt des QQNorm-Plots und der Dichtefunktion sind im unteren Teil der Grafik erkennbar. Der QQNorm-Plot wird durch die Standardisierung nur in Bezug auf die Lage der Daten verändert. Dies ist an der y-Achse erkennbar. An der Gestalt des QQNorm-Plots ändert sich nichts. Dies ist nicht überraschend, da die Standardisierung der Daten nichts an dem Verhältnis der theoretischen Quantile der Normalverteilung und der Quantile der Daten zueinander verändert.

Anhand der Dichtefunktionen ist erkennbar, dass sich die Daten durch die Standardisierung hinsichtlich ihrer Lage und Streuung verändern. Die Dichtefunktion der Standardnormalverteilung ist in der Grafik unten rechts durch die gestrichelte Linie eingezeichnet. Auf Basis dieser Grafik kann dem Studierenden der Sinn der Transformationen der Daten verdeutlicht werden. Die Transformationen führen zu einer Annäherung der Gestalt der Dichtefunktion an die der Standardnormalverteilung. Für diese existieren Tabellenwerte, die zum statistischen Testen verwendet werden. In der traditionellen Statistikausbildung wurden die Tabellenwerte zum Testen verwendet. Durch die Umsetzungen im Statistiklabor wird deutlich, warum die Tabellenwerte auf die Daten angewendet werden können.

Die folgende Statistiklaborseite zeigt, dass die Logarithmierung als erste Transformation deutlich besser geeignet ist als zum Beispiel das Ziehen der 2. Wurzel:

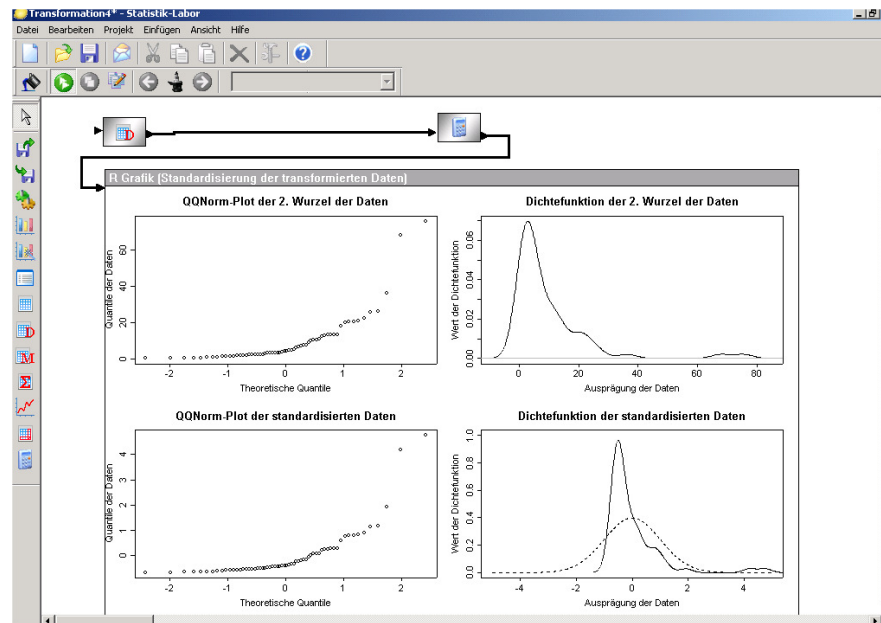


Abbildung 23: Interaktivität: Standardisierung der transformierten Daten

Die Notwendigkeit der Wahl einer geeigneten Transformation, in diesem Fall die Überlegenheit der Logarithmierung, wird auch durch den direkten Vergleich deutlich:

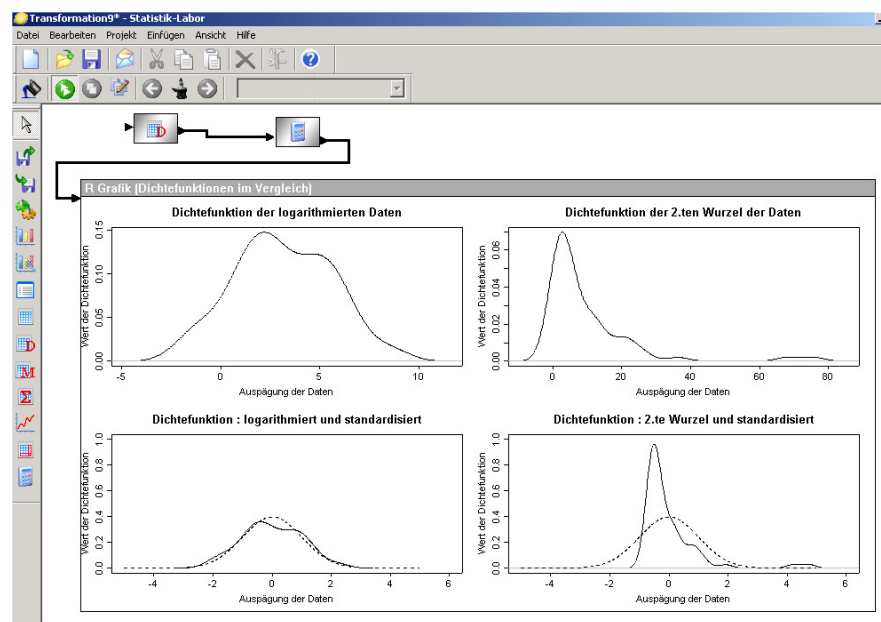


Abbildung 24: Vergleich zweier Transformationen mit anschließender Standardisierung

Abschließend kann der gesamte Transformationsprozess auf einer Statistiklaborseite zusammenfassend dargestellt werden. Es besteht außerdem die Möglichkeit der Interaktivität. Der Studierende kann den Prozess mit unterschiedlichen Arten der Transformation durchlaufen. Diese Möglichkeit und der Umgang mit der Laborseite sind wie oben in einem Texteditor aufgeführt.

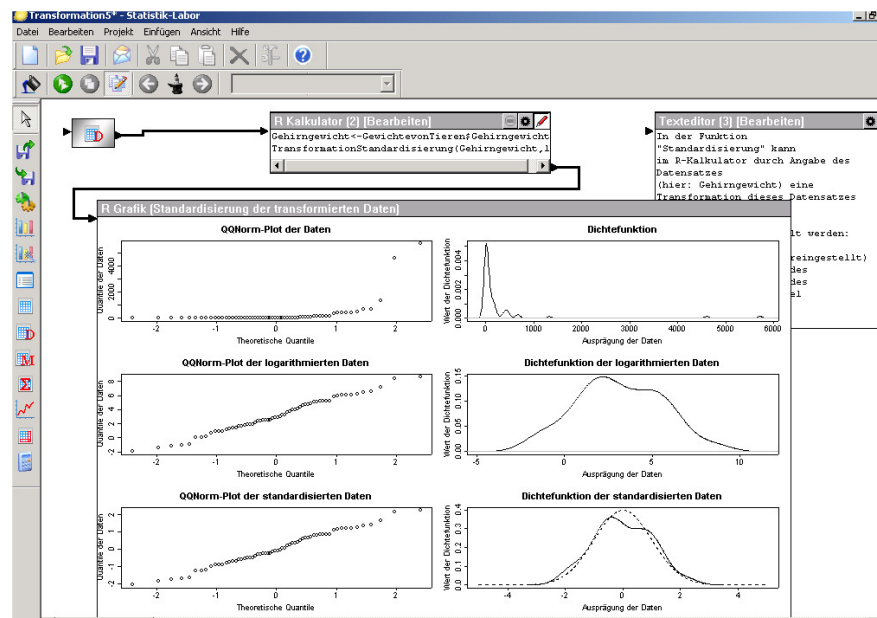


Abbildung 25: Interaktivität: Der gesamte Transformationsprozess

Durch die in diesem Abschnitt demonstrierten Statistiklaborseiten kann dem Studierenden die Auswirkung von Transformationsprozessen vermittelt werden. Er kann unterschiedliche Transformationsprozesse interaktiv auf Datensätze anwenden. Er kann mit den beiden bereitgestellten oder mit eigenen Datensätzen experimentieren.

**Alternatives Datenmaterial** Um zu verdeutlichen, dass die Logarithmierung nicht in jedem Fall die optimale Transformation ist, Daten in eine normalverteilte Gestalt umzuwandeln, wird ein weiterer Datensatz in die Diskussion eingebracht. Es handelt sich um *Dart Scores*:

”Christine enjoys playing darts and conducted a study to investigate her ability. Christine’s study consisted of 40 trials. A trial consists of Christine throwing three darts at a standard dart board, and the response is the total score for the three darts.”<sup>121</sup>

Auf der folgenden Statistiklaborseite wird demonstriert, dass für den vorliegenden Datensatz die Transformation *Ziehung der zweiten Wurzel* besser geeignet ist, als die Logarithmierung:

<sup>121</sup>[Wardrop2003].

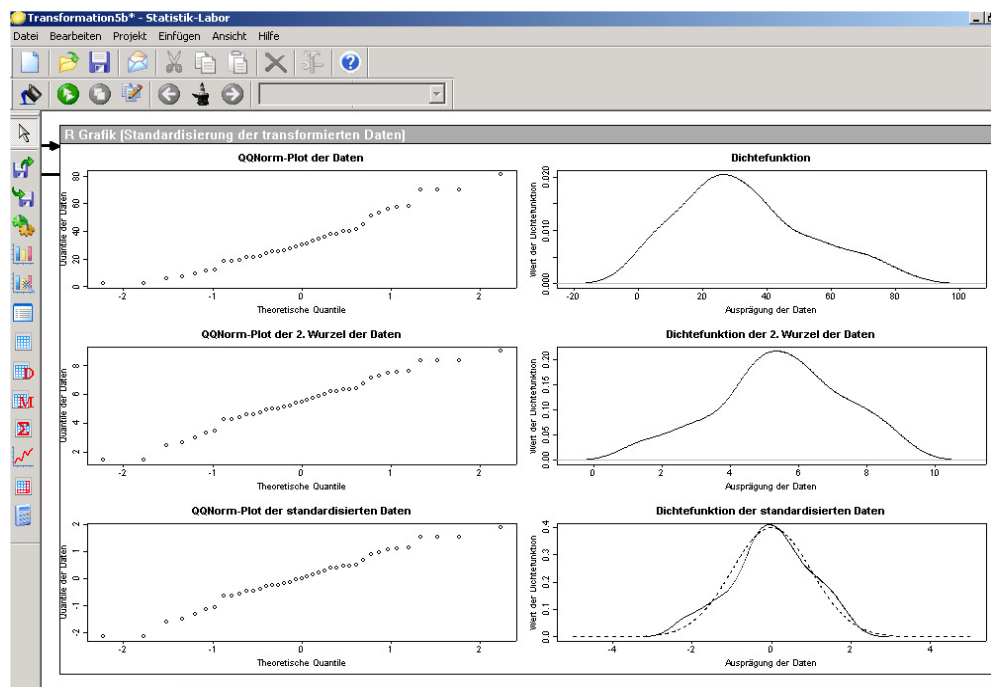


Abbildung 26: Transformation: Ziehung 2ter Wurzel mit anschließender Standardisierung

Es besteht die Möglichkeit der Interaktivität - wählt man die Transformation *Logarithmierung*, so entstehen die folgenden Grafiken:

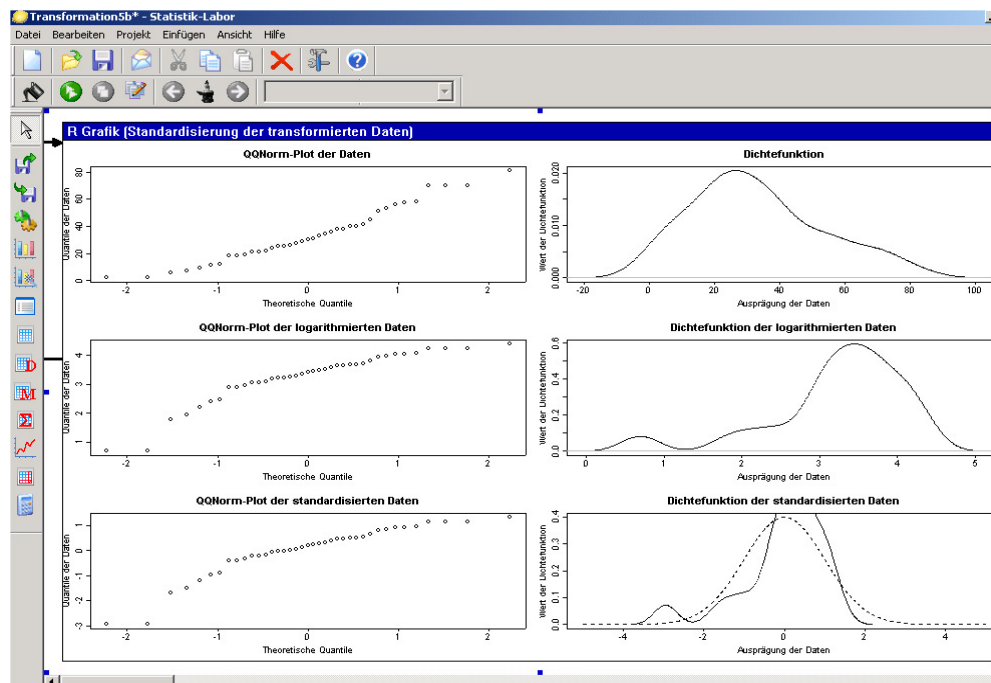


Abbildung 27: Transformation: Logarithmierung mit anschließender Standardisierung

Die bessere Eignung der Transformation *Ziehung der zweiten Wurzel* ist eindeutig erkennbar.

**Zum Einsatz in der Statistikausbildung** Bezogen auf die Standardisierung könnten in einer Theoriekomponente die folgenden Hintergründe erläutert werden:

- In einer Grundstudiumveranstaltung könnte erklärt werden, dass die Standardisierung  $z = \frac{x-\mu}{\sigma}$  sowohl die Lage (Erwartungswert) als auch die Variabilität der Daten (Standardabweichung) verändert.
- Eine weitere Theoriekomponente könnte erklären, dass die nach der Transformation vorliegenden Daten nur unter bestimmten Bedingungen annähernd normalverteilt sind: Unter den bisherigen Annahmen sind die Daten t-verteilt. Die t-Verteilung kann allerdings für Stichprobenumfänge größer 30 durch die Normalverteilung approximiert werden.<sup>122</sup>

**Bezug zur Defizitklassifikation** Zusammenfassend kann festgestellt werden, dass die neuen Medien einen besseren Transport der statistischen Inhalte zu ermöglichen scheinen, da sie Interaktivität des Studierenden hervorrufen. Die Prinzipien des Konstruktivismus (*active learner*) und die Möglichkeiten der neuen Medien wurden kombiniert eingesetzt, um Defizite der Statistikausbildung entgegenzuwirken. Betrachtet man abschließend die im Rahmen dieser Arbeit aufgestellte Defizitklassifikation und begründen, welche Defizite konkret angesprochen werden:

- Dem Defizit B (fehlende Interaktivität/Aktivität der Studierenden) wird durch die entwickelten Statistiklaborseiten entgegengewirkt. Der Studierende kann zwischen verschiedenen Transformationsprozessen wählen und die ihre Auswirkungen und Unterschiede grafisch kennenlernen. Er hat die Möglichkeit, experimentell zu lernen und sich im Sinne des Konstruktivismus sein Wissen durch Eigenaktivität zu konstruieren. Die verschiedenen Laborseiten ermöglichen, dass dieser Prozess schrittweise abläuft, d.h. altes mit neuem Wissen nach konstruktivistischem Prinzip miteinander verknüpft wird.
- Da die Transformationsprozesse mit realen Daten durchgeführt werden, wird dem Defizit C (fehlender Realitätsbezug) entgegengewirkt.
- Durch den Einsatz des R-Kalkulators/des Statistiklabors werden die Möglichkeiten des Rechners demonstriert - Computational Statistics wird vermittelt (Defizit G). An dieser Stelle kann der Dozent entscheiden, wie weit er ins Detail geht. Die obigen Statistiklaborseiten enthalten im R-Kalkulator die Funktionsaufrufe. Ebenso kann der Dozent die dahinterstehenden Bibliotheken aufrufen und ihre einzelne Schritte diskutieren, je nachdem wie weit er die Studierenden an das Programmieren heranzuführen möchte. Wieder kommt der Konstruktivismus zum Tragen. Die Funktionalität des R-Kalkulators kann sowohl durch die Vernetzung von vorhandenem mit neuem Wissen erlernt werden als auch durch eigene Aktivität im Statistiklabor.

---

<sup>122</sup>Vgl. [Johnson1992], S. 101.



#### 4.4 Problem der geeigneten Stichprobengröße

**Daten, Problem, Idee, Defizit** Den Ausgangspunkt bildet der Datensatz *Fisher's Irises*, der in der Vergangenheit bereits für viele Datenanalysen und Forschungen herangezogen wurde, deshalb auch als ein *famous dataset* bezeichnet wird. Fisher erhob und benutzte die Daten für die Illustration der Diskriminanzanalyse. Der Datensatz enthält fünf Variablen mit je 150 Beobachtungen. Fisher hat 150 Schwertlilien vermessen. Diese enthalten drei verschiedene Gruppen von Schwertlilien: Setosa, Verginica und Versicolor. Von jeder Gruppe sind genau fünfzig Schwertlilien in dem Datensatz enthalten. Fisher hat an jeder Schwertlilie die folgenden vier Merkmale erhoben:

1. Petal Width - Weite des Blütenblattes
2. Petal Length - Länge des Blütenblattes
3. Sepal Width - Weite des Kelchblattes
4. Sepal Length - Länge des Kelchblattes

Das Ziel dieses Kapitels besteht darin, mit Hilfe des Statistiklabors Probleme der Statistik zu demonstrieren. Die Vorteile des Rechners, speziell des Statistiklabors für die Ausbildung sollen hervorgehoben werden.

Es sollen Informationen über eine Grundgesamtheit gewonnen werden. Für diesen Zweck hat eine Datenerhebung stattgefunden. Es handelt sich um eine Teilerhebung. Anhand dieser werden die Probleme der Stichprobenziehung vermittelt. Hierbei wird gezeigt, welche Auswirkungen eine Verkleinerung des Stichprobenumfangs haben kann.

Anhand der einzelnen Gruppen von Schwertlilien wird das Problem der geeigneten Stichprobengröße diskutiert. Es wird der Frage nachgegangen, wie groß eine Stichprobe gewählt werden muss, um repräsentative Aussagen über die Grundgesamtheit treffen zu können. Dabei wird sowohl über die Größe der Stichprobe, als auch die jeweilige Größe der in ihr enthaltenen Gruppen argumentiert. In dem betrachteten Datensatz enthält jede der drei Gruppen genau 50 Ausprägungen, d.h. Fisher untersuchte 50 Setosa-, 50 Verginica- und 50 Versicolorschwertlilien. Diese Gruppengrößen lassen eine Argumentation über das Gesetz der großen Zahlen zu. Das Ziel besteht darin, dem Studierenden folgende Aspekte und ihre Zusammenhänge zu verdeutlichen:

- Anhand von Fisher's Irises Datensatz wird die Idee der Stichprobenziehung vermittelt. Es wird verdeutlicht, dass der Stichprobenumfang sowie die Anzahl der Beobachtungen innerhalb der Gruppen in der Stichprobe eine entscheidende Rolle spielen. In diesem Zusammenhang wird das Gesetz der großen Zahlen empirisch beobachtet und theoretisch erklärt.
- Der Bezug zur Varianzzerlegung wird hergestellt.<sup>123</sup> Die Voraussetzungen und ihre Anwendung werden anhand der erhobenen Daten vermittelt. Dies geschieht insbesondere über die drei Gruppen von Schwertlilien.

---

<sup>123</sup>Es folgt eine etwas andere Herangehensweise, als in dem bereits vorgestellten Lösungsvorschlag, um den Studierenden auf verschiedenen Wegen an die Varianzzerlegung heranzuführen: oben wurde ihre Idee, jetzt die Voraussetzungen und Anwendungen diskutiert.

- Während der Diskussion der voranstehenden Aspekte wird der Zusammenhang zwischen der Empirie und der Theorie bei statistischen Überlegungen und Analysen herausgestellt. Die Empirie wird durch die Datenerhebung widerspiegelt, die Theorie durch das Gesetz der großen Zahlen.

Um diese Aspekte zu vermitteln, wird das Statistiklabor der multimedialen Lernsoftware *Statistik interaktiv komplett* eingesetzt. Die konkreten Umsetzungen werden im folgenden Unterkapitel vorgestellt und diskutiert. Die Entwicklung von Statistiklaborseiten hat das Ziel, die oben genannten statistischen Inhalte und Probleme besser zu transportieren als die bisherige Statistikausbildung.

**Statistiklaborseiten - Demonstration und Interaktivität** Das Defizit in der traditionellen Statistikausbildung besteht darin, dass große Datenmengen nicht während der Vorlesung analysiert werden können. Aus diesen Gründen gestaltete sich die Vermittlung der Stichprobenziehungen und der damit verbundenen Probleme als schwierig. Betrachtet man die Ideen zur Vermittlung der oben genannten Aspekte auf den folgenden Statistiklaborseiten. Diese sollen demonstrieren, dass die neuen Medien konstruktivistische Vermittlungsmöglichkeiten und Prinzipien in die Statistikausbildung bringen können.

**Laborseite 1:** Der Datensatz *Fisher's Irises* wird im Detail betrachtet. Es werden die Dichtefunktionen der vier erhobenen Merkmale (Weite des Blütenblattes, Länge des Blütenblattes, Weite des Kelchblattes und Länge des Kelchblattes) im Statistiklabor grafisch dargestellt. Betrachten wir die Laborseite<sup>124</sup>:

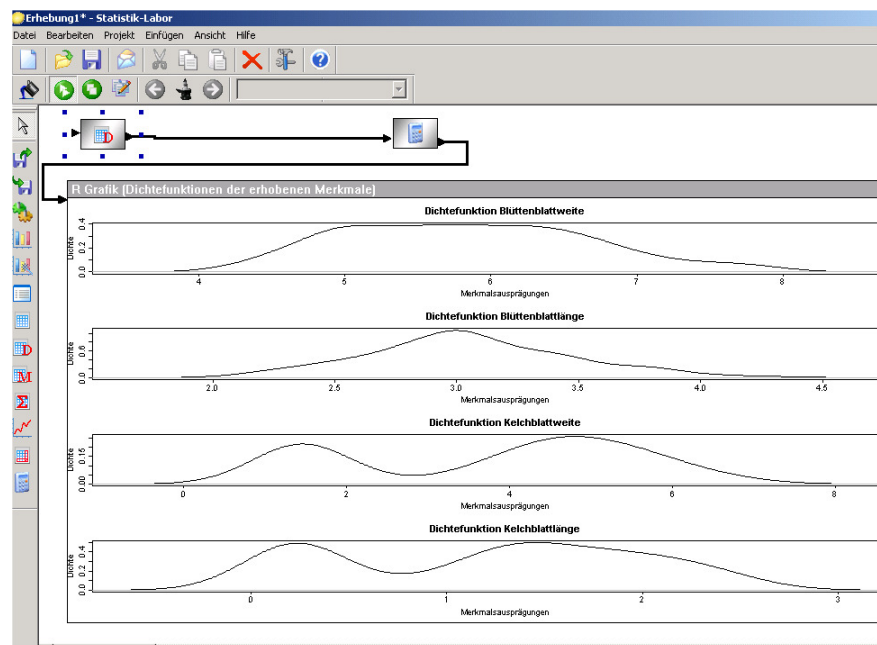


Abbildung 28: Dichtefunktionen der vier erhobenen Merkmale

<sup>124</sup>Diese Laborseite eignet sich zu Demonstrationszwecken, beinhaltet keine Interaktionsmöglichkeiten.

Obwohl der Datensatz drei Gruppen von Schwertlilien enthält, lassen die Dichtefunktionen der Merkmale nicht auf einen trimodalen Charakter der Daten schließen. Die beiden unteren Dichtefunktionen lassen eher auf einen bimodalen Charakter schließen. Es scheint zu einer Überlagerung der Dichtefunktionen zu kommen. Diesen Aspekt betrachtet man auf der folgenden Laborseite genauer.

Technische Bemerkung: Man wird beobachtet haben, dass die unteren beiden Dichtefunktionen Merkmalsausprägungen im negativen Bereich besitzen. Inhaltlich ständen diese für negative Längen und Breiten der Kelchblätter. Die Ursache liegt in der angewendeten R-Funktion *density()*: die Dichtefunktionen werden aus den empirischen Daten geschätzt (Dichteschätzer). Es besteht die Möglichkeit, bei dem Dichteschätzer die Fensterbreite zu verringern, so dass eine sehr große Anpassung der Dichtefunktion an die empirischen Daten stattfindet. Problematisch ist hierbei zum einen die grafische Darstellung:

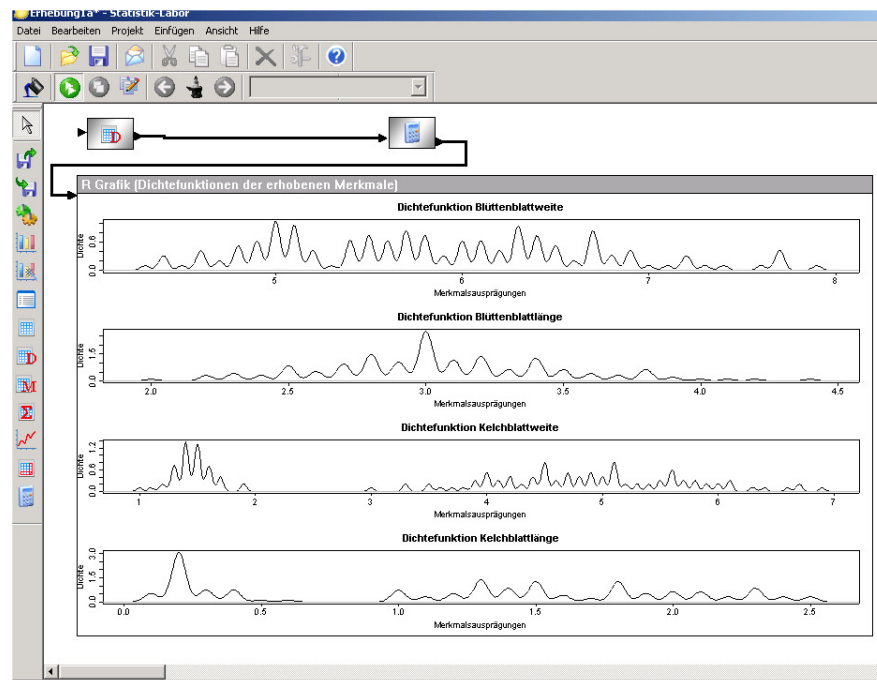


Abbildung 29: Dichtefunktion mit veränderter Fensterbreite

Zum anderen besteht die Problematik der zu großen Anpassung des Datensatzes (overfitting), so dass der Dichteschätzer nur für diesen Datensatz geeignet ist. Aus diesem Grund ist die erste Grafik gewählt, um die Dichtefunktion darzustellen.

**Laborseite 2:** Auf dieser Statistiklaborseite werden die Dichtefunktionen der drei Gruppen von Schwertlilien (*Setosa*, *Verginica*, *Versicolor*) betrachtet:

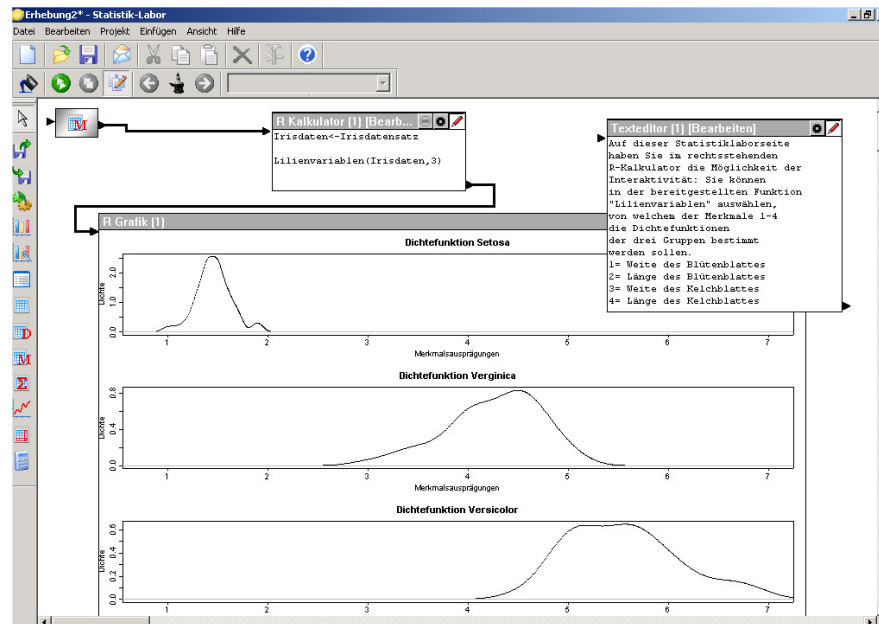


Abbildung 30: Dichtefunktionen der drei Gruppen

Es besteht die Möglichkeit der Interaktivität. Der Studierende kann auswählen, für welches der vier erhobenen Merkmale er sich die Dichtefunktionen der Gruppen anzeigen lassen möchte. Die Auswahl wird ihm in dem Texteditor erklärt und findet im R-Kalkulator in der bereitgestellten Funktion *Lilienvariablen* statt.

Voreingestellt ist das Merkmal *Weite des Kelchblattes*. Die zugehörigen Dichtefunktionen lassen erkennen, warum auf der Laborseite 1 trotz dreier Gruppen nur ein bimodaler Charakter der Daten erkennbar war. Die Dichtefunktion der Gruppen *Verginica* und *Versicolor* überlagern sich.

**Laborseite 3:** Anhand dieser Statistiklaborseite soll die Problematik der Wahl der minimalen Stichprobengröße demonstriert werden. Das arithmetische Mittel des Gesamtdatensatzes der Iris Daten und die Annäherung des Datensatzes bei steigendem Stichprobenumfang werden grafisch dargestellt. Es wird die Botschaft transportiert, dass erst bei hinreichend großem Stichprobenumfang eine Annäherung an das arithmetische Gesamtmittel stattfindet. Die Grafik enthält vertikale Linien. An ihnen werden zwei Werte berechnet und angegeben: der Stichprobenumfang und das entsprechende arithmetische Mittel. Der Studierende kann auf diese Weise zusätzlich zur Grafik numerisch erkennen, dass bei steigendem Stichprobenumfang eine bessere Anpassung an das arithmetische Mittel des Gesamtdatensatzes stattfindet.

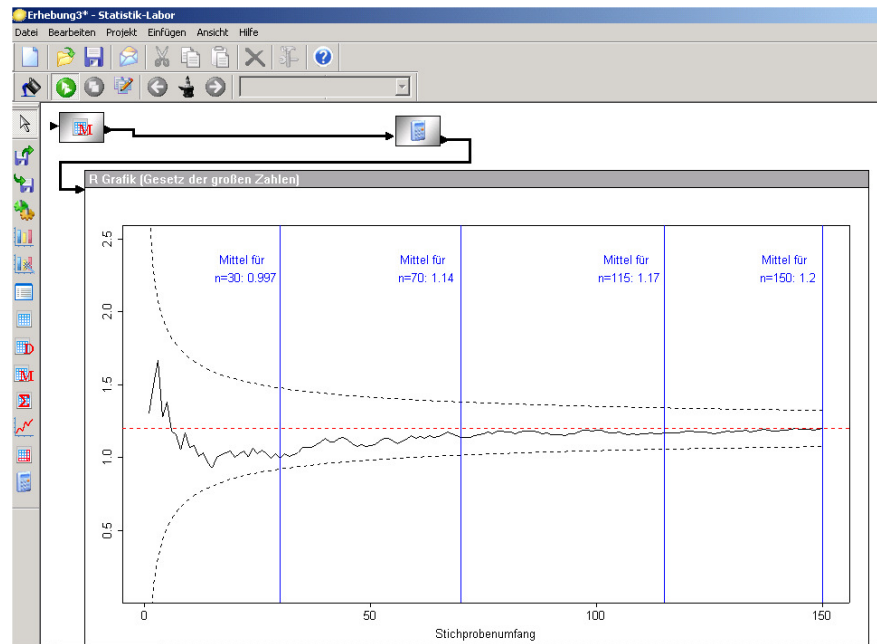


Abbildung 31: Demonstration des Gesetzes der großen Zahlen

An dieser Stelle kann der Zusammenhang zwischen Empirie und Theorie hergestellt werden. Das Gesetz der großen Zahlen (Theorie) kann in Verbindung mit dem realen Datensatzes (Empirie) erklärt werden. Eine **Theoriekomponente** könnte direkt mit der Statistiklaborseite verknüpft werden. Des Weiteren kann zu dem in der Lernsoftware *Statistik interaktiv komplett* enthaltenen Lernmodul *Erhebungsverfahren* verknüpft werden. Dieses vermittelt die Grundlagen der Stichprobenziehung.

The screenshot shows a Microsoft Internet Explorer browser window. The address bar contains the URL: [http://bb-portal.cedis.fu-berlin.de/courses/1/PS-UBI-WIWI-Statistik1\\_035/content/\\_4127\\_1/dir\\_1\\_ZIP/1/lermodule/03\\_Erhebungsverfahren/03\\_02\\_Ein](http://bb-portal.cedis.fu-berlin.de/courses/1/PS-UBI-WIWI-Statistik1_035/content/_4127_1/dir_1_ZIP/1/lermodule/03_Erhebungsverfahren/03_02_Ein). The page title is "Warum Stichproben?". The main content area is titled "Grundlagen der Stichprobenziehung bei Bevölkerungsbefragungen" and includes a dropdown menu with "Warum Stichproben?". The text explains that sampling is used because it is more cost-effective and faster than a full census. It lists several reasons for using sampling, such as cost, time, and the risk of damage. A sidebar on the right contains three sections: "MEDIEN" (Animation, Inferenzpopulation, Animation, Inferenzschluss), "ÜBUNGEN" (Wie würden Sie entscheiden: Stichprobe oder Vollerhebung?, Vollerhebung und Stichprobe, Voraussetzungen für den Inferenzschluss, Vorteile von Stichproben in der Praxis, "overcoverage" und "undercoverage"), and "BEISPIELE" (Vorteile einer Stichprobe, Under- und Overcoverage bei Bevölkerungstichproben, Ansoestrebte).

Abbildung 32: Lernmodul Erhebungsverfahren

Die Umsetzung im Statistiklabor ist in diesem Lernmodul nicht enthalten, könnte aber durch die Ausführungen dieses Kapitels sinnvoll ergänzt werden.

**Laborseite 4:** Aus dem Gesamtdatensatz der 150 Schwertlilien werden zufällig Vollerhebungen gezogen. Es wird immer aus Fisher's Irisdatensatz gezogen, die Reihenfolge ist einem Zufallsprozess überlassen. Der Studierende kann in der im R-Kalkulator bereitgestellten Funktion *irisauswahl* auswählen, wie viele Stichproben gezogen werden sollen.

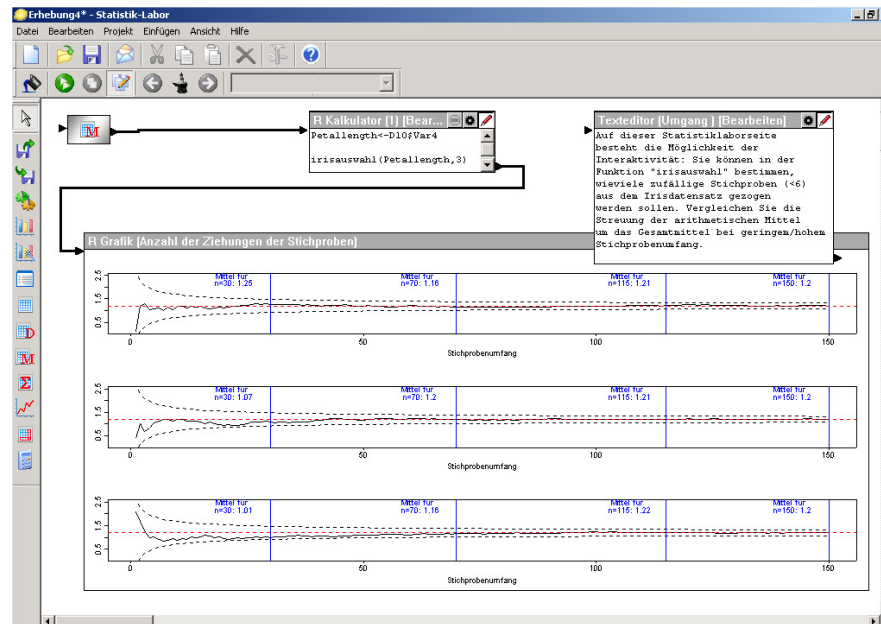


Abbildung 33: Verschiedene Stichprobenziehungen aus Fishers Irisdaten

Der Vergleich der verschiedenen Anordnungen zeigt zum einen, dass die Werte des arithmetischen Mittels bei geringem Stichprobenumfang sehr stark um das arithmetische Gesamtittel des Irisdatensatzes streuen, während bei hohem Stichprobenumfang eine viel geringere Streuung vorhanden ist. Durch die Möglichkeit der Interaktion kann der Studierende ein Gefühl für die Wahl der Stichprobengröße bekommen. Hier wird zunächst nur die Ziehung einer Stichprobe betrachtet. Auf den Laborseiten 6 und 7 werden zusätzlich die Gruppen innerhalb dieser (Setosa, Verginica, Versicolor) berücksichtigt.

**Laborseite 5:** Nach der Ziehung zufälliger Stichproben aus dem Gesamtdatensatz auf der Laborseite 4 wird nun die Möglichkeit gegeben, in Bezug auf den Gesamtdatensatz die Stichprobengröße zu variieren. Zu dieser werden jeweils die arithmetischen Mittelwerte in der Grafik angezeigt. Der Texteditor erklärt den Umgang mit der Statistiklaborseite und lenkt den Blick des Studierenden:

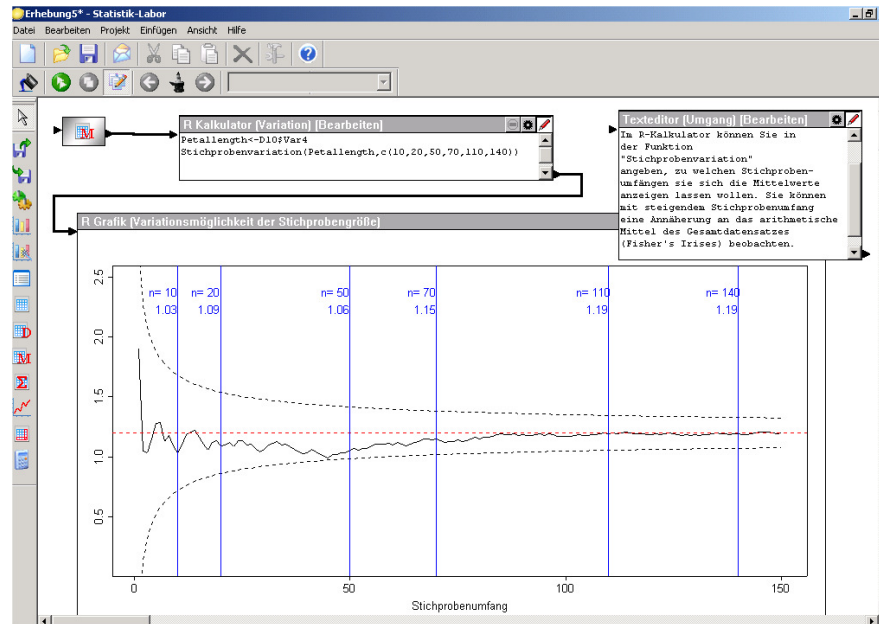


Abbildung 34: Variation der Stichprobenumfänge mit angezeigten Mittelwerten

**Laborseite 6:** Auf dieser Statistiklaborseite wird auf Basis des Gesetzes der großen Zahlen vermittelt, warum nicht nur die Größe des Stichprobenumfangs des Gesamtdatensatzes (Fisher's Irises), sondern auch die Stichprobenumfänge der in ihm enthaltenen Gruppen (Setosa, Verginica, Versicolor) entscheidend für die Repräsentativität der Stichprobe sind. Betrachten wir zunächst die Laborseite, die keine Interaktivität ermöglicht, sondern Demonstrationszwecken dient:



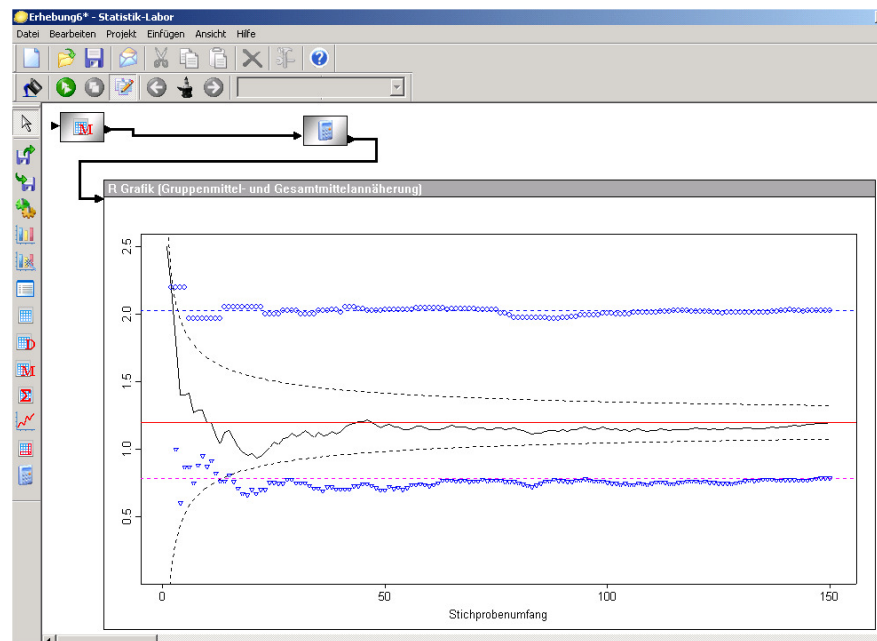


Abbildung 35: Gesetz der großen Zahlen - Gruppenmittel und Gesamtmittel

Die trichterförmigen Sigtabänder unterstreichen das stochastische Konvergieren des Verlaufs der Mittelwerte des Gesamtdatensatzes bei steigendem Stichprobenumfang gegen sein arithmetisches Mittel. Zusätzlich sind die Verläufe der Gruppenmittelwerte grafisch dargestellt: oberhalb des Gesamtdatensatzes die der Gruppe *Setosa*, unterhalb die der Gruppen *Verginica* und *Versicolor*. Die beiden Schwertliliengruppen *Verginica* und *Versicolor* werden zusammengefasst, da sich der Verlauf ihrer Dichtefunktionen überlagert (Vgl. Statistiklaborseite 1). In dieser Grafik würde die getrennte Betrachtung beider Gruppen zu Unübersichtlichkeit führen. Die Daten liegen zu nah beieinander.

Es ist deutlich erkennbar, dass sich die arithmetischen Mittel der einzelnen Gruppen auch erst mit steigendem Stichprobenumfang an das jeweilige Gruppenmittel des gesamten Irisdatensatzes annähern. Der Verlauf des aus *Verginica* und *Versicolor* zusammengesetzten Datensatzes liegt näher an dem des Gesamtdatensatzes, da doppelt so viele Daten in ihm enthalten sind, wie in dem der Gruppe *Setosa*.

An dieser Stelle kann der Bogen zur Varianzzerlegung geschlagen werden. Dem Studierenden kann verdeutlicht werden, dass die Identifikation von Gruppen in Datensätzen zur Erklärung der Streuung entscheidend sein kann. Diese Erklärung enthält eine etwas andere Blickrichtung auf das Problem der Varianzzerlegung als der voranstehende Lösungsvorschlag.<sup>125</sup> Es wird über die sich bei steigendem Stichprobenumfang bildenden Mittelwerte argumentiert anstatt über die Varianz. Die Annäherung der Gruppen von Schwertlilien an ihre arithmetischen Mittel und die Beeinflussung des Verlaufes des Gesamtdatensatzes wird auf der Laborseite 6 deutlich. Bei der Varianzzerlegung wird die Gesamtstreuung zerlegt in die Streuung zwischen den Gruppen und innerhalb der Gruppen. In der Grafik ist die Streuung zwischen den Gruppen 1 (*Setosa*) und 2 (*Verginica*+*Versicolor*) deutlich er-

<sup>125</sup>Vgl. 4.1

kennbar. Die Streuung innerhalb der beiden Gruppen ist bei geringem Stichprobenumfang grafisch abgebildet. Die einzelnen Mittelwerte streuen um das jeweilige Gruppenmittel.

**Laborseite 7:** Diese Statistiklaborseite baut auf der Laborseite 6 auf und ermöglicht Interaktivität. Der Studierende kann angeben, zu welchen Stichprobenumfängen die Datensätze der Gruppen sowie des zusammengesetzten Datensatzes abgebildet werden sollen. Er kann diese miteinander vergleichen und durch Experimentieren feststellen, ab welchem Stichprobenumfang sowohl in den Gruppen als auch im zusammengesetzten Datensatz eine gute Annäherung an ihre arithmetischen Mittelwerte stattfindet. Die Gruppenmittel wie auch der Gesamtmittelwert werden in jedem Fall auf Basis des Stichprobenumfangs 150 gebildet.

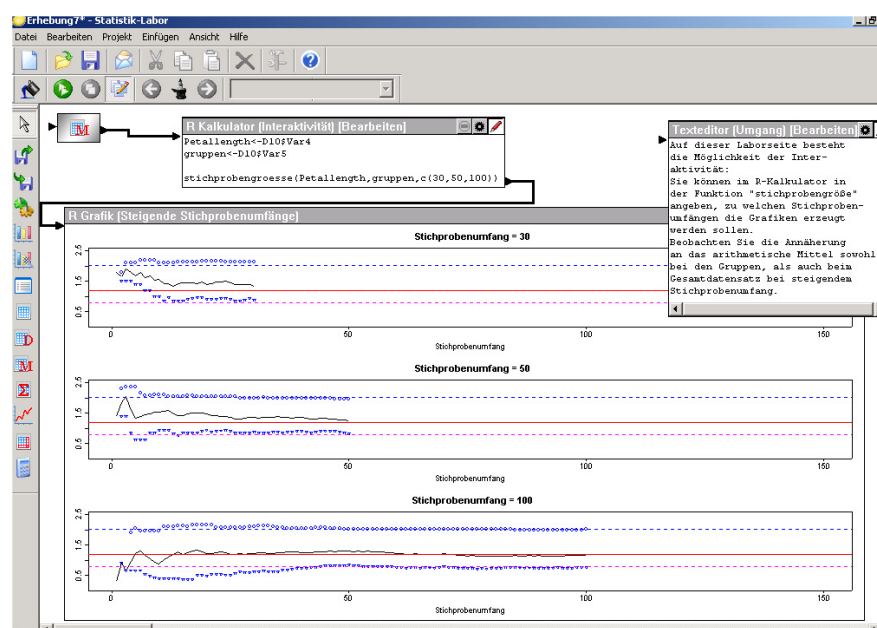


Abbildung 36: Interaktivität: Gruppenmittel und Gesamtmittel

Es ist deutlich zu erkennen, dass bei einem Stichprobenumfang von 30 eine deutlich schlechtere Annäherung der Gruppendaten als auch des Gesamtdatensatzes an die jeweiligen Mittelwerte stattfindet. Mit Hilfe dieser Statistiklaborseite kann erklärt werden, warum das Gesetz der großen Zahlen erst ab einem Stichprobenumfang von dreißig gilt.

Die Statistiklaborseite ermöglicht konstruktivistischen Statistikerunterricht. Auf Basis der grafischen Darstellungen kann gemeinsam diskutiert werden, warum die Stichprobengröße eine entscheidende Rolle für das Gesetz der großen Zahlen spielt, warum Methoden/Gesetze nicht uneingeschränkt Anwendung finden können.

**Bezug zur Defizitklassifikation** Zusammenfassend kann festgestellt werden, dass die neuen Medien einen besseren Transport der statistischen Inhalte ermöglichen, da sie Interaktivität des Studierenden hervorrufen. Betrachten wir abschließend die im Rahmen dieser Arbeit aufgestellte Defizitklassifikation und begründen, welchen Defiziten durch die Umsetzung der präsentierten Lösungsvorschläge positiv entgegenwirkt werden kann:

- Dem Defizit B (fehlende Interaktivität/Aktivität der Studierenden) wird durch die entwickelten Statistiklaborseiten entgegengewirkt. Der Studierende kann Stichprobenumfänge auswählen, die Anzahl von zu ziehenden Stichproben verändern und zwischen Merkmalen des Datensatzes wählen. Er hat die Möglichkeit, experimentell zu verstehen, dass die Größe der Stichproben einen Einfluss auf die Repräsentativität der Stichprobe hat.
- Durch die Argumentation auf Basis des realen Datensatzes *Fisher's Irises*, wird dem Defizit C (fehlender Realitätsbezug) entgegengewirkt.
- Durch den Einsatz des R-Kalkulators/des Statistiklabors werden die Möglichkeiten des Rechners demonstriert, Computational Statistics wird vermittelt (Defizit G). An dieser Stelle kann der Dozent entscheiden, wie weit er ins Detail geht. Die obigen Statistiklaborseiten enthalten im R.Kalkulator die Funktionsaufrufe. Ebenso kann der Dozent die dahinter stehenden Bibliotheken aufrufen und einzelne Schritte diskutieren, je nachdem wie weit er die Studierenden an das Programmieren heranführen möchte.

**Ausblick:** Es ist geplant, die Lernsoftware *Statistik interaktiv komplett* in Zukunft zusammen mit multimedialen Lösungen dieser Art bereitzustellen.

## Teil II

# Evaluation multimedialer Lernsoftware - das Fundament

**Überblick** Im voranstehenden Teil I wurde die Statistik charakterisiert. Die Rolle und Entwicklung der Statistikausbildung wurden skizziert. Im Zentrum stand die Identifikation der in der Statistikausbildung existierenden Defizite. Diese wurden in Defizitgruppen unterteilt und in einer Defizitklassifikation zueinander in Beziehung gesetzt. Anschließend wurden Lösungsvorschläge entwickelt und umgesetzt. Identifizierten Defiziten wurde durch den Einsatz des Statistiklabors entgegengewirkt. Durch die Lösungsvorschläge wurden die Möglichkeiten der neuen Medien, insbesondere des Statistiklabors, für die Neugestaltung der Statistikausbildung herausgestellt.

Das Ziel dieses Teils besteht in der Beantwortung der Frage, ob Multimedia bzw. die neuen Medien eine Chance oder Alternative für die Statistikausbildung darstellen. Es wird zunächst diskutiert, was unter den vielseitig verwendeten Begrifflichkeiten *Multimedia* und *Neue Medien* verstanden wird. Anschließend wird der Versuch unternommen, den Begriff *Evaluation* zu definieren. Drei allgemeine Ansätze zur Evaluation werden herangezogen, um einen Raum aufzuspannen. Durch diesen und die kritische Diskussion der Ansätze wird die Position des Autors in Bezug auf Evaluation herausgestellt. Anschließend wird gezeigt, welche entscheidende Rolle Evaluation im Zusammenhang mit den neuen Medien spielt. Die facettenreiche Diskussion um die *Evaluation von Multimedia* wird betrachtet. Die Darstellung verschiedener Ansätze verdeutlicht dem Leser den Wandel und die neue Rolle des Evaluationsbegriffes. Es werden Konsequenzen für die in Teil IV dargestellte Evaluationsstudie abgeleitet.

**Einschränkungen und Abgrenzungen** Die vorliegende Arbeit bewegt sich in einem Themengebiet (*Multimedia, neue Medien* und *Evaluation*), das aus verschiedenen Perspektiven und Disziplinen betrachtet wird. Eine Abgrenzung im Voraus ist daher notwendig. Folgende Bereiche werden nicht behandelt, da sie Multimedia zwar betreffen, aber in keinem direkten Zusammenhang zu den Schwerpunkten dieser Arbeit stehen:

- Im Rahmen der Entwicklung wird nicht auf Hardware-Konfigurationen, spezielle Betriebssysteme, Systemschnittstellen oder technische Spezifikationen eingegangen.
- Es werden keine Ausführungen zu Themen wie Tele-Teaching, Video-on-Demand, Digital Audio und Video, Video-Conferencing vorgenommen, die häufig mit Multimedia in Verbindung gesetzt werden.
- Das Online-Lernen wird im Rahmen dieser Arbeit nicht näher erläutert.
- Es wird keine Auflistung und Klassifikation aller möglichen Typen von Multimedia-Anwendungen vorgenommen.
- Das Konzept der *Virtuellen Universität* wird in dieser Arbeit nicht aufgegriffen, da es sich um ein eigenes, in sich komplexes wie auch umstrittenes Gebiet handelt, das nicht unmittelbar im Zusammenhang mit den Zielen dieser Arbeit steht.

Die folgenden Fragen werden untersucht:

- Was versteht man unter Multimedia/Neuen Medien?
- Verändert sich durch die neuen Medien die Qualität der (Statistik-)Ausbildung?
- Was bedeutet Evaluation von Multimedia/Neuen Medien?

## 5 Multimediale Lernsoftware - Definitionsversuche und Reflexionen

### Definitionsversuche

„Multimedia is the seamless integration of data, text, images of all kinds and sound within a single, digital information environment.“<sup>126</sup>

Es existiert eine Vielzahl von Definitionsversuchen auf diesem Gebiet. Begriffe wie *Multimedia*, *Neue Medien* oder *Neue Statistik*, die noch vor fünf Jahren in keinem Wörterbuch zu finden waren, begegnen uns heute in den unterschiedlichsten Bereichen und Disziplinen.

„Interdisziplinäre Verständigungsschwierigkeiten sind in der Diskussion um Multimedia vorprogrammiert.“<sup>127</sup>

Ein Blick in die neuere Literatur zeigt, dass diese Aussage eine realistische Situationsbeschreibung ist. Die unzähligen Versuche, die Begriffe *Multimedia* oder *Neue Medien* zu definieren, zeigen, dass es sich bei Multimedia um etwas Vielseitiges und somit schwer Greifbares handelt. Ob Informatik, Psychologie, Didaktik, Pädagogik oder Wirtschaftswissenschaften - jede Disziplin scheint etwas anderes unter Multimedia zu verstehen. Jeder rückt andere Aspekte in den Mittelpunkt seiner Überlegungen. Betrachten wir exemplarisch einige Definitionsversuche, um dem Leser einen Einblick in die Situation zu geben:

„Multimedia ist ein Konzept, das nicht nur die Medien, sondern auch die technische und die anwendungsbezogene Dimension integriert.“<sup>128</sup>

„Multimedia ist nicht nur eine relativ neue Technologie mit wachsender wirtschaftlicher Bedeutung. Multimedia ist auch eine neue Art und Weise der Mediennutzung in Informations- und Lernprozessen. Multimedia wird wie alle neue Medien durch die Eigenschaften der Individualität, Interaktivität, Asynchronität und Multifunktionalität charakterisiert.“<sup>129</sup>

„Unter Multimedia versteht man also die gleichzeitige oder zeitlich versetzte Verwendung mehrerer statischer und dynamischer Medien auf einer Präsentationsplattform, (...)“<sup>130</sup>

---

<sup>126</sup>Feldmann in [Schulmeister2002], S. 16.

<sup>127</sup>[IssingKlimsma1997], S. 7.

<sup>128</sup>[IssingKlimsma1997], S. 9.

<sup>129</sup>[IssingKlimsma1997], S. 1.

<sup>130</sup>[IssingKlimsma1997].

„Multimedia bezeichnet das Zusammenspiel aller derzeit verfügbaren elektronischen Datenträger für Bild- und Toninformationen in der Information, Aus- und Weiterbildung und Unterhaltung. Der Anwender ist aktiv beteiligt und kann den Ablauf nach seinen Wünschen gestalten.“<sup>131</sup>

Am Ende dieses Abschnittes wird ein eigener Definitionsversuch für Multimedia vorgenommen. Dabei werden Aspekte dieser Zitate, wie z.B. die Rolle des Anwenders und seine Interaktivität mit dem System aufgegriffen. Die Erfahrungen, die im Rahmen der Evaluationsstudie<sup>132</sup> in Bezug auf multimediale Lernsoftware gesammelt wurden, haben gezeigt, welche Eigenschaften den Lernerfolg entscheidend beeinflussen.

Betrachten wir ein Zitat aus einer Mitteilung des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie aus dem Jahr 1998:

„Die Wissensgesellschaft ist die Vision des 21. Jahrhunderts. Nur ihre Umsetzung kann im Zeitalter zunehmender Globalisierung den Standort Deutschland langfristig sichern. Die Wege zu wirtschaftlichem Wachstum und mehr Beschäftigung führen über die Entwicklung und Verbreitung der neuen Informations- und Kommunikationstechnologien. Der Umgang mit ihnen muss zur Selbstverständlichkeit werden. (...) Für die aktive Gestaltung des Weges in die Wissensgesellschaft ist das erklärte Ziel des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie (BMBF): Multimedia möglich machen.“<sup>133</sup>

Diese Worte spiegeln hohe (eventuell zu hohe) Erwartungen an den Einsatz von Multimedia in der Bildung wider, ohne dass erklärt wird, was unter Multimedia zu verstehen ist und was der Satz *Multimedia möglich machen* bedeuten soll.

Eine mögliche Auslegung des Satzes könnte lauten: Die Entwicklung qualitativ hochwertiger Multimediasoftware hängt von deren Einsatz ab. Dieser muss konzipiert werden, um zielgerichtet zu erfolgen. Die vorliegende Arbeit stellt die These auf, dass der Einsatz multimedialer Lernsoftware auf Basis eines speziell entwickelten Einsatzkonzeptes basieren sollte. Ein Einsatzkonzept legt den Umgang mit den Komponenten der Lernsoftware fest, die zielgerichtet miteinander kombiniert werden. In Teil IV wird die Lernsoftware *Statistik interaktiv komplett* auf Basis eines Einsatzkonzeptes evaluiert. Im Einsatzkonzept werden verschiedene Komponenten der entwickelten Lernsoftware *Statistik interaktiv komplett* so kombiniert, dass sie identifizierten Defiziten der Statistikausbildung<sup>134</sup> entgegenwirken können.

„Mit Multimedia auf der Datenautobahn in die Informationsgesellschaft. Stärker läßt sich die derzeitige (politische) Programmatik wohl nicht mehr verknäppeln. Multimedia steht in dieser Kombination für „medienreiche Dokumente“, die Datenautobahn für „Hochgeschwindigkeitsnetze“ und die Informationsgesellschaft für das Gesamtsystem aus technischer Infrastruktur, Kommunikationsmodell und Selbstverständnis dieser Gesellschaft.“<sup>135</sup>

Mit dieser Aussage leiten Riehm/Wingert ihre Äußerungen zu Chancen und Herausforderungen von Multimedia ein. Technisch definiert sich Multimedia für sie über die Kombination diskreter und kontinuierlicher Medien. Für deren Zusammenspiel und Interaktion

---

<sup>131</sup>[Börner1992].

<sup>132</sup>Vgl. 14.2.

<sup>133</sup>[BMBF1998].

<sup>134</sup>Vgl. 4.4.

<sup>135</sup>[RiehmWingert1995], S. 1.

(oder „interaktive Nutzbarkeit“) ist der Computer unverzichtbar. Das heißt, dass Multimedia in einer ersten Näherung aus den Komponenten Computer, Medien und Interaktion besteht.

„Im Kern geht es heute bei Multimedia um die Interaktion mit computerbasierten Anwendungen, in denen unterschiedliche Medientypen integriert sind. (...) Multimedia ist keine neue Technologie an sich, sondern die Zusammenführung bisher getrennter Technologien und Anwendungen.“<sup>136</sup>

Sie warnen vor einer Überschätzung der neuen Möglichkeiten und betonen, entgegen aller intuitiven Logik, dass mehr Medien mehr bringen<sup>137</sup>, zu beachten: Das Lernen mit Multimedia ist nur unter bestimmten Bedingungen effektiv, die erst herausgefunden werden müssen.

„Die mediale Form für sich bewirkt es nicht; es kommt ganz entscheidend auf die methodische und didaktische Aufbereitung und die richtige Einbettung in den Lernkontext an.“<sup>138</sup>

Zwei Erkenntnisse können abgeleitet werden:

- Die Einbettung von Multimedia in den Lernkontext erfordert die Entwicklung von Einsatzkonzepten.
- Die didaktische Aufbereitung erfordert den Einsatz einer geeigneten Lerntheorie. Dem Leser wurde dargelegt, dass konstruktivistische Prinzipien bei der Entwicklung von *Statistik interaktiv komplett* und des Einsatzkonzeptes eine zentrale Rolle spielen.

**Definitionsversuch dieser Arbeit** Die bisherigen Ausführungen haben demonstriert, dass die Diskussion um Multimedia weitläufig und damit schwer greifbar ist. Die folgende Aussage fasst die Situation zusammen:

„Die uneinheitliche Verwendung des Begriffs *Multimedia*, die verschwimmenden Grenzen zu anderen Informations- und Kommunikationstechniken, die Fülle von Anwendungen, die unter dem Begriff subsumiert werden - alles dieses macht deutlich, dass Multimedia eine Chiffre darstellt, unter der sich vieles verbirgt, und die im öffentlichen Wissen noch keine Kontur gewonnen hat.“<sup>139</sup>

Es wird deshalb nicht danach gestrebt, *die* allgemeingültige Begriffsdefinition zu finden. Der Autor nimmt für diese Arbeit den folgenden Definitionsversuch vor:

**„Die neuen Medien/Multimedia<sup>140</sup> ermöglichen das Zusammenwirken unterschiedlicher (klassischer und multimedialer) Medien in einem System. Die Gestaltung der neuen Medien stellt die Interaktion des Anwenders mit diesem System in den Mittelpunkt. Mögliche Interaktionen sind der Austausch von Informationen und das Experimentieren. Die anwenderbezogene Orientierung sowie die Interaktion des Anwenders auf einem für ihn durch verknüpfte Medien bereitgestellten Experimentierfeld hat das Ziel, den Lernprozess voranzutreiben.“**

<sup>136</sup>[RiehmWingert1995], S. 8.

<sup>137</sup>Vgl. [RiehmWingert1995], S. 249.

<sup>138</sup>[RiehmWingert1995], S. 4f.

<sup>139</sup>[Kindt1999], S. 240.

<sup>140</sup>Die Begriffe werden synonym verwendet.

Die Schwierigkeit der Begriffsdefinition wird deutlich. Auch dieser Definitionsversuch ist weit auslegbar. Er legt nicht fest, wie die unterschiedlichen Medien zusammenwirken oder was unter einem System verstanden wird. Dennoch werden zentrale Aspekte, die im Rahmen dieser Arbeit bei neuen Medien eine Rolle spielen, festgelegt: das Zusammenwirken, der Anwender, die Interaktion, das Experimentieren.

Die Arbeit basiert auf der These, dass die neuen Medien neue Möglichkeiten für die Statistikausbildung darstellen können, wenn sie zielgerichtet entwickelt und eingesetzt werden. Baumgartner stärkt diese These, indem er zum Einsatz von Multimedia in der Lehre bemerkt:

„Wir glauben, dass mit Multimedia die Qualität der Lehre unter bestimmten Bedingungen, unter bestimmten Voraussetzungen verbessert werden kann. Und bitte nicht umgekehrt, wie: zuerst Multimedia, und dann schauen wir mal, was wir mit der Qualität der Lehre machen können.“<sup>141</sup>

Ein großer Vorteil von Multimedia besteht darin, Lernprozesse individueller gestalten zu können, auch heterogene Gruppen lehren zu können und verschiedene Angebote mit unterschiedlichen Zielsetzungen nutzen zu können. Dem ersten Teil des folgenden Zitats schließt sich der Autor an:

„Vom sporadischen Aufsuchen individuell interessanter Informationen bis zur Absolvierung eines ganzen Kursprogrammes mit Prüfungen - anything goes!“<sup>142</sup>

Das euphorische *anything goes* ist kritisch zu betrachten. Jede Studie, jedes Multimedia- oder Evaluationsprojekt der letzten Jahre berichtet über Probleme, die durch Multimedia entstehen. Die folgende Reflexion über die neuen Medien in der Lehre gibt dem Leser einen Einblick in die Problematik.

**Neue Medien in der Lehre - eine Reflexion** Sowohl Gestaltungs- und Entwicklungsfragen als auch die Erfüllung von Erwartungen seitens der Nutzer werden kritisch beleuchtet und diskutiert. Es wird differenziert der Frage nachgegangen, ob die Multimediaentwicklung auf dem richtigen Weg ist. Über bisherige Anstrengungen, Aktivitäten, Multimediaprojekte wird reflektiert. Es wird eine kritische Bilanz gezogen.

„Multimedia in der Lehre - ist ein Schlagwort, das je nach Perspektive als didaktische Innovation, organisatorische Revolution oder finanzpolitische Sparmaßnahme mit effizienzsteigernder Wirkung verstanden werden kann.“<sup>143</sup>

„Mit der Diskussion um die Einführung und Verbreitung neuer, computer- und netzbasierter Medien an den Hochschulen ist implizit und explizit die Erwartung auf Veränderung eines gegebenen Zustandes verbunden. Ob es sich dabei um eine Verbesserung oder Verschlechterung handelt, hängt vom Standpunkt der Betrachtung ab.“<sup>144</sup>

Beide Zitate sind kennzeichnend für die Diskussion um den Einsatz von Multimedia in der Lehre. Ein Standpunkt, der selten eine Rolle spielt, ist der der finanziellen Förderer. Ihr

---

<sup>141</sup>[Kindt1999], S. 134.

<sup>142</sup>[Kindt1999], S. 63-99.

<sup>143</sup>[Hauff1998], S. 11.

<sup>144</sup>[Hauff1998], S. 17.



vorrangiges Ziel besteht darin, Defizite des derzeitigen Hochschulsystems zu kompensieren. Sie vertreten die Ansicht, dass Medienkompetenz eine wichtige Schlüsselqualifikation auf dem zukünftigen Arbeitsmarkt darstellen wird. Aus diesem Grund sollen folgende Lernformen an den Hochschulen verstärkt gefördert werden: individuelles Lernen, interaktives Lernen, kooperatives Lernen, interdisziplinäres Lernen, globales und dynamisches Lernen. Dabei fehlt es größtenteils nicht an technischen Entwicklungen multimedialer Software, sondern vielmehr an flexiblen Organisationsstrukturen und Einstellungen.

Lömker dazu:

„Es fehlt nicht an individuellen Initiativen. Es fehlt aber an profilbildenden Hochschulstrategien. (...) Multimediaunterstützung für die Lehre lässt sich in verschiedenen Stufen denken. Zum Teil sind damit erhebliche Umstrukturierungen des Studiums und der Hochschule verbunden.“<sup>145</sup>

Die in Teil IV dieser Arbeit vorgestellte Evaluationsstudie kommt zu dem Ergebnis, dass die zur Zeit noch starren Strukturen der Hochschulen den Einsatz neuer Medien in der Lehre erschweren, innovative Lehr- und Lernmaßnahmen blockieren und sie mit sehr hohem Aufwand versehen.

In seinem Artikel *Multimedia in der Hochschullehre* äußert sich Keil-Slawik zu den sogenannten „Infrastrukturnotwendigkeiten“:

„Eine weltweite multimediale Informationsstruktur mit neuen Diensten entsteht. Diese durch Digitalisierung ermöglichte Verschmelzung von Medien verkörpert die spezifische neue Qualität von Multimedia und hat weitreichende Konsequenzen für die Produktion, Verteilung und Erschließung von Inhalten, denn ohne die Einbettung der Medien in entsprechende organisatorische und institutionelle Abläufe der Produktion, Selektion, Aufbewahrung, Verknüpfung und Erschließung bleiben sie so wertlos wie eine unentzifferte Hieroglyphenschrift.“<sup>146</sup>

Solche Infrastrukturüberlegungen erscheinen insbesondere im Rahmen der Hochschule notwendig. Keil-Slawik verdeutlicht, dass Wechselwirkungen zwischen folgenden Größen bestehen:

- der Didaktik und der Aufbereitung der Materialien seitens der Dozenten,
- der Entwicklung und dem Einsatz der neuen Medien,
- der Technik bzw. den Medien, die zum Einsatz kommen sollen,
- den Institutionen (hier: Hochschulen)

Er berichtet über den Versuch, sich im Rahmen der Bertelsmann-Stiftung durch eine Art Bestandsaufnahme einen Überblick über den Multimediaeinsatz an deutschen Universitäten zu verschaffen. Das Ziel dieses Vorhabens bestand darin, Aussagen darüber treffen zu können, welche qualitativen Verbesserungen durch Multimedia in der Lehre erreicht werden konnten und unter welchen Bedingungen. Es wurde der Versuch unternommen, 28 Projekte zu vergleichen. Das Ergebnis der qualitativen Auswertungen hatte nicht die erwünschte Aussagekraft. Die auftretenden Probleme bei dem Vergleich waren der Meinung der Autorin nach aus den folgenden Gründen vorhersehbar:

---

<sup>145</sup>[Hauff1998], S. 19.

<sup>146</sup>[Hauff1998], S. 29ff.

- Viele Projekte befanden sich erst in der Entwicklungsphase.
- Es lagen keine konzeptionierten Evaluations- und Einsatzkonzepte vor.
- Viele Projekte waren so speziell, dass keine wiederholte Nutzung der Materialien, somit keine nutzerspezifische schrittweise Verbesserung stattfinden konnte.
- Es konnten keine qualitätsrelevanten Faktoren identifiziert werden, da zuvor keine Ziele definiert wurden.

Der Hauptgrund für die enttäuschenden Ergebnisse ist laut Keil-Slawik der, dass es bei einer Vielzahl von Projekten kein einheitliches Verständnis von Multimedia gab. Außerdem war ein Vergleich der Projekte nicht möglich, da sie sich zu stark bezüglich Konsequenz, Umfang und Qualität unterschieden. Die Erfahrungen des Einsatzes von Multimedia in der Statistikausbildung unterstützen diese Aussagen. Der didaktische, organisierte und in lerntheoretische Konzepte eingebettete Multimediaeinsatz ist so individuell, dass bei einem Vergleich verschiedener Projekte Schwierigkeiten aufgrund fehlender Abgrenzungskriterien auftreten. Es scheint keine Möglichkeit zu geben, einen effizienten Vergleich der auf dem Hochschulmarkt befindlichen Multimediaprojekte durchzuführen. Zu dem jetzigen Zeitpunkt kann man festhalten, dass es keine Kriterien gibt, die eine klare Abgrenzung aller Multimediaeinsätze in der Lehre und eine Bewertung des aktuellen Standes ermöglichen. Keil-Slawik bezeichnet diese Tatsache als *Evaluationsdefizit* von Multimedia und führt es zurück auf die Anfangs- und Experimentierphase, in der sich ein Großteil der Projekte zur Zeit noch befindet. Keil-Slawik beschreibt ein weiteres Defizit:

„Es handelt sich um Initiativen einzelner Forscherpersönlichkeiten, die mit viel Aufwand und einer hohen Motivation die erforderlichen technischen und organisatorischen Voraussetzungen schaffen. (...) Mit der Beendigung des Projektes oder dem Weggang der Initiatoren werden die diesbezüglichen Aktivitäten an der jeweiligen Hochschule eingestellt.“<sup>147</sup>

Betrachtet man einige Multimediaprojekte der letzten Jahre und bezieht die eigenen Erfahrungen mit ein, so kann diese Situation auf den hohen Aufwand für Entwicklung und Einsatz von Multimedia in den Lehrveranstaltungen zurückgeführt werden. Umfangreiche Maßnahmen sind erforderlich, um von traditioneller auf multimediale Lehre umzustellen. Die Erfahrungen, die im Forschungsprojekt *Neue Statistik* gesammelt wurden, zeigen, dass keine vollständige Umstellung von traditioneller auf multimediale Lehre benötigt wird. Einsatzkonzepte für multimediale Lernsoftware haben insbesondere die Aufgabe, die Instrumentarien beider zu kombinieren. Bei der Konzeption des Einsatzes multimedialer Lernsoftware wäre es ideal, wenn der erhöhte Aufwand zu einem höheren Nutzen führte. Dies kann dann der Fall sein, wenn zum Beispiel dynamische Abläufe oder spezifische Formen der Interaktion, experimentelle Erzeugung von Daten oder Visualisierungen komplexer dynamischer Zusammenhänge ermöglicht werden. Dies sind Eigenschaften, die von den traditionellen Medien nicht dargestellt werden können, oder durch die Neuen Medien besser erfasst werden können.

Keil-Slawik vertritt die Meinung, dass für eine dauerhafte, effektive Erschließung der neuen Qualitäten von Multimedia umfangreiche didaktische, organisatorische und curriculare Umstrukturierungen erforderlich sind, um die verschiedenen Medienwelten zu integrieren. Er behauptet, dass sich Multimedia erst da lohnt, wo traditionelle Verfahren versagen

---

<sup>147</sup>[Hauff1998], S. 29ff.

oder ineffektiv sind. Der Leser wird in der im Teil IV dieser Arbeit<sup>148</sup> dokumentierten Evaluationsstudie erkennen, dass multimediale Lernsoftware auch eingesetzt werden kann, um traditionelle Instrumentarien zu ergänzen.

Keil-Slawik stützt seine Argumentation bezüglich der Infrastrukturänderungen auf die These, dass man mit Technik nur technische Probleme lösen kann, Lehren und Lernen aber vorrangig mit sozialen und pädagogischen Problemen zu tun haben.<sup>149</sup> Es werden lernförderliche Infrastrukturen benötigt, um diese beiden Bereiche miteinander zu verbinden. Betrachtet man die zahlreichen Diskussionen der letzten Jahre, so fallen zwei Extrema auf. Befürworter neigen zu einer unrealistischen Überbetonung der neuen Medien, der neuen technischen Möglichkeiten (je mehr Technik, desto besser der Lernprozess). Die andere Position nehmen diejenigen ein, die die neuen Medien und neuen Techniken ablehnen und von einem Verlust des Lernens durch Multimedia sprechen. Bei beiden wird deutlich, dass Technik und Inhalt gemeinsam betrachtet werden müssen.

Diese Arbeit ergänzt die Ausführungen durch die Integration des Konstruktivismus. Sowohl die Technik als auch die Inhalte müssen nach konstruktivistischen Prinzipien gestaltet sein. Die Lösungsvorschläge und exemplarischen Umsetzungen des Teils I<sup>150</sup> haben unter anderem das Zusammenwirken von statistischen Inhalten, neuen Medien und konstruktivistischen Prinzipien demonstriert.

Die zahlreichen rund um die Entwicklung, Gestaltung, Evaluation und den Einsatz neuer Medien geführten Diskussionen setzen verschiedene Schwerpunkte. Folgende Aspekte sind im Rahmen dieser Arbeit von Bedeutung:

- Neue Medien verlangen neue Gestaltungen neuer Informations- und Kommunikationssysteme.
- Neue Medien verlangen eine neue Art der didaktischen Aufbereitung der Inhalte.
- Neue Medien verlangen neue Organisationsstrukturen in der Hochschule.
- Neue Medien werfen alte Fragen der Didaktik und Lerntheorien in einem neuen Zusammenhang wieder auf. Baumgartner bemerkt, dass es wichtig ist, dass Lehr- und Prüfungsdidaktik übereinstimmen. Wird mediengestützt gelehrt und gelernt, muss die Prüfungsform diesen neuen Medien und dem neuen medialen Lernen angepasst werden und darf nicht in traditionellen Strukturen verharren.

Baumgartner<sup>151</sup> bezeichnet die neuen Medien als einen Katalysator einer anderen Sichtweise und äußert sich im Zusammenhang mit der Konstruktion neuer Lehr- und Lernwelten über die von Keil-Slawik gewählte Metapher *Multimedia als Steinbruch des Lernens*:

„Für interessant halte ich jedoch die durch diese Metapher angeregte Vorstellung, dass nicht jeder irgendwo, irgendwie, irgendwann und ohne jede Absprache mit anderen „Bauherren“ etwas herausbrechen kann. Gerade auch ein Steinbruch erfordert eine gemeinsame Planung und Organisation und Sachkenntnis über die vorhandene Unter- bzw. Grundlage. Die Metapher lenkt das Bild auf

---

<sup>148</sup>Vgl. 14.2.

<sup>149</sup>Vgl. [Hauff1998], S. 81ff.

<sup>150</sup>Vgl. 4.

<sup>151</sup>Vgl. [Hauff1998], S. 57ff.

das organisatorische Gesamtsystem, weg von den einzelnen Steinblöcken hin auf den gesamten Steinbruch.“<sup>152</sup>

Diese Gedanken werden im Teil III<sup>153</sup> bei der Entwicklung der Lernsoftware *Statistik interaktiv komplett* aufgegriffen. Es wird das Ziel verfolgt, die Lernsoftware modular zu gestalten. Die einzelnen Module sollen zusammengefügt ein in sich schlüssiges System bilden. Es wird diskutiert, mit welchen Problemen die modulare Gestaltung verbunden ist. Bezüglich der inhaltlichen Gestaltung der Lernmodule und Komponenten wird ein weiterer Aspekt berücksichtigt: die Rolle von Lerntheorien. Im Folgenden wird der Zusammenhang zwischen neuen Medien und traditionellen Lerntheorien erläutert.

### Neue Medien und Theorien des Lernens - der Konstruktivismus

„Aus heutiger Sicht ist es sicher nicht sinnvoll, für jede neu entdeckte Art der Wissensaneignung eine eigene Lernform zu etablieren. (...) Die Hoffnung liegt darin, mit Multimedia gerade die Formen des einsichtigen und komplexen Lernens (z.B. des Problemlösens) zu unterstützen.“<sup>154</sup>

Die konstruktivistische Lerntheorie scheint am besten geeignet, den klassifizierten Defiziten in der Statistikausbildung entgegenzuwirken. Es wurde gezeigt, dass die Statistikausbildung erste konstruktivistische Elemente enthält. Dieser Trend soll durch die Entwicklung und den konzipierten Einsatz von *Statistik interaktiv komplett* verstärkt werden. Durch die Integration konstruktivistischer Elemente in die Lernmodule und das Einsatzkonzept wird den klassifizierten Defiziten der Statistikausbildung entgegengewirkt. Dabei wird sich zeigen, dass der Konstruktivismus genau wie die neuen Medien kein Allheilmittel zur Defizitbekämpfung darstellt. Es wird demonstriert, wie konstruktivistische Ideen und Elemente in der multimedialen Lernsoftware *Statistik interaktiv komplett* und dem speziell entwickelten Einsatzkonzept verankert sind.

„ (...) the optimal starting point for understanding the constructivist perspective to teaching and learning is to consider what constructivism is not. Constructivism is often articulated in stark contrast to the behaviorist model of learning.“<sup>155</sup>

“The role of education is to help students learn about the real world. The goal of designers or teachers is to interpret events for them.“<sup>156</sup>

Dem ersten Teil der Aussage schließen wir uns an, vermitteln statistische Methoden anhand realer Datensätze. Die Rolle des Dozenten sehen wir allerdings darin, die Studierenden beim Konstruieren von Interpretationen zu unterstützen. Das Ziel der Lehre besteht zudem nicht ausschließlich darin, Konzepte zu vermitteln, sondern auch eine positive Bewertung durch den Studierenden hervorzurufen.<sup>157</sup> Eine Bewertung setzt die schrittweise Entwicklung eines individuellen Bewertungssystems beim Lernenden voraus: Er entscheidet mit Hilfe des Dozenten, ob er neue Konzepte für richtig befindet, sie als gut bewertet.

<sup>152</sup>[Hauff1998], S. 59.

<sup>153</sup>Vgl. 10.

<sup>154</sup>[Hauff1998], S. 155f.

<sup>155</sup>[Murphy1997].

<sup>156</sup>Jonassen in [Murphy1997].

<sup>157</sup>Dem Leser ist dieser Aspekt bereits in der Differenzierung zwischen P- und Mstatistics begegnet.

Die Evaluationsstudie zeigt, dass zum Lernen neuer Konzepte ein Ungleichgewicht gehört. Lernen wird dadurch ausgelöst, dass sich der Studierende nicht immer sicher über sein existentes, konstruiertes Wissen ist. Er überdenkt bestehendes Wissen und Werte.

Dieser konstruktivistischen Aspekt wurden bei der Entwicklung von *Statistik interaktiv komplett* berücksichtigt. Ein Ungleichgewicht wird in verschiedenen Lernmodulen und Komponenten ausgelöst. Die Musterreports demonstrieren dem Benutzer ein Problem und verschiedene Lösungswege. Auch in den Fallstudien werden verschiedene Perspektiven bei der Betrachtung konkreter Probleme eingenommen. Bei der Erstellung der eigenen Reports ist die Problemstellung jeweils so formuliert, dass die Studierenden zwischen der Anwendung verschiedener statistischer Instrumente wählen können. Glasersfeld bemerkt zu diesem Aspekt:

“It is not the particular response that matters but the way in which it was arrived at.”<sup>158</sup>

“Having constructed a viable path of action, a viable solution to an experiential problem, or a viable interpretation of a piece of language, there is never any reason to believe that this construction is the only one possible.”<sup>159</sup>

Multimedia bietet Möglichkeiten, diese Idee des Konstruktivismus neu und vielleicht besser als bisher umzusetzen. Studierende neigen dazu, einen Lösungsweg als den einzig richtigen zu betrachten, wenn er einige Male problemlos funktioniert hat. Sie verlieren dabei aus den Augen, dass es nur einer vieler Wege ist, die zum Ziel führen. Insbesondere bei der Vermittlung der statistischen Reporterstellung lernen die Studierenden, während der Lösung ihrer Probleme verschiedene Wege zu erforschen. Sie geben sich nicht mit der erstbesten Lösungsmöglichkeit zufrieden, sondern suchen nach Alternativen.

Der Konstruktivismus spielte bei der Gestaltung multimedialer Lernsoftware eine Rolle. Wie konstruktivistische Ideen konkret umgesetzt wurden, ist zum einen in den in Teil I präsentierten Lösungsvorschlägen deutlich geworden. Zum anderen werden konkrete Beispiele in den folgenden Teilen demonstriert. Verschiedene Komponenten der Lernsoftware und des Einsatzkonzeptes werden detailliert im Hinblick auf die Anwendung konstruktivistischer Ideen erläutert. Es findet zudem eine Bewertung von *Statistik interaktiv komplett* und des Einsatzkonzeptes auf Basis einer konstruktivistischen Checkliste statt.

**Zwischenfazit:** Nach der Beschreibung verschiedener Facetten des Multimediabegriffes wurde ein eigener Definitionsversuch unternommen, der dieser Arbeit zugrunde liegt. Dem Leser wurde Einblick in die Diskussion um die neuen Medien in der Lehre gegeben. Das Ziel bestand darin, Erkenntnisse für den Einsatz multimedialer Lernsoftware in der Statistikausbildung abzuleiten. Viele Aussagen konnten durch eigene Erfahrungen untermauert werden, die im Rahmen der Evaluationsstudie gesammelt wurden. Die eigentliche Frage, die es hier zu beantworten gilt, ist die nach der Form des Einsatzes von Multimedia in der Statistikausbildung. Wie soll Multimedia eingesetzt werden, um die Statistikausbildung zu bereichern, d.h. bestehenden Defiziten entgegenzuwirken? Dieses *wie* setzt sich aus zwei Gebieten zusammen: der Entwicklung<sup>160</sup> und dem Einsatz<sup>161</sup> multimedialer Lernsoftware.

---

<sup>158</sup>[Glasersfeld1995].

<sup>159</sup>[Glasersfeld1995].

<sup>160</sup>Vgl. 10.

<sup>161</sup>Vgl. 14.2.

In beiden Bereichen müssen bestimmte Voraussetzungen geschaffen werden, beide müssen kombiniert betrachtet werden, um die Qualität der Lehre mit Hilfe von Multimedia verbessern zu können. Die Arbeit stellt die These auf, dass es keine natürliche Reihenfolge dieser Gebiete gibt, sondern dass es notwendig ist, iterativ vorzugehen, Erkenntnisse auf dem einen Gebiet in das andere einfließen zu lassen, Zwischenergebnisse des einen auf dem anderen zu erproben. Im Teil III wird auf Basis dieser Erkenntnis ein Vorgehensmodell zu der Entwicklung und dem Einsatz der multimedialen Lernsoftware *Statistik interaktiv komplett* entwickelt. Wichtig ist die Abstimmung beider Gebiete. Keil-Slawik sagt in diesem Zusammenhang:

„In Bezug auf Multimedia ist davon auszugehen, dass die Erfolgchancen in hohem Maße von Integration und Abstimmung abhängig sind. Integration bedeutet, dass Technik, Didaktik und curriculare Entwicklung nicht isoliert betrachtet werden dürfen: Neue Qualitäten ergeben sich erst, wenn alle Komponenten gleichermaßen berücksichtigt werden. Die besondere Herausforderung besteht darin, dass schon das Versagen an einer einzigen Stelle, wie z.B. mangelnde Verfügbarkeit der technischen Ressourcen oder fehlende Qualifikationen auf Seiten der Lehrenden, den Nutzen in Frage stellt.“<sup>162</sup>

Die konkreten Umsetzungen der Gestaltungsaspekte werden im Teil III aufgegriffen. Zuvor wird die Evaluation von Multimedia charakterisiert. Es werden Konsequenzen für die explorative Evaluationsstudie abgeleitet.

## 6 Evaluation - verschiedene Perspektiven

Die neuen Chancen und die heillose Überschätzung von Multimedia in der Bildung führten dazu, dass im Laufe der letzten Jahre seitens des Bundes viel Geld in sogenannte *Multimedialprojekte* geflossen ist. Die Ergebnisse ließen zu wünschen übrig. Um diesem Phänomen entgegenzuwirken, stellte das Bundesministerium den Anspruch der Evaluation an die Multimedialprojekte.

„Die Forderung, dass es für die Ausgabe von Milliarden von Mark nicht ausreichend ist, zu sagen, wir machen etwas Gutes, sondern die effektive Verwendung der Mittel transparent gemacht werden muss, ist legitim und nachvollziehbar.“<sup>163</sup>

Ein Blick in die Literatur zeigt, dass diese Aussage Keil-Slawiks<sup>164</sup> unter Wissenschaftlern verbreitet ist.

Es werden allerdings Stimmen laut, die betonen, dass es seit eh und je zum wissenschaftlichen Arbeiten gehört hat, seine Ergebnisse zu überprüfen. So Keil-Slawik auf dem Abschlusspodium der Tagung zum Thema *Projektelevaluation in der Lehre*:

„Wissenschaftliches Ziel eines Projektes ist, dass ich etwas lerne. Um etwas zu lernen, brauche ich Feedback, ich muss sozusagen Rückmeldung haben. Da fange ich doch jetzt nicht plötzlich an, über Evaluation von Projekten zu reden. Wenn ein Projekt ein Forschungsprojekt ist, dann ist es natürlicher Bestandteil, Evaluation zu machen. Ich nenne es bloß nicht so, denn es geht um das

---

<sup>162</sup>[Hauff1998], S. 40.

<sup>163</sup>[Kindt1999], S. 128f.

<sup>164</sup>Professor an der Universität-GH Paderborn, Fachgebiet: Informatik und Gesellschaft

klassische wissenschaftliche Instrumentarium, die eigenen Ergebnisse zu überprüfen, an Zielen zu messen und zu hinterfragen, was man verallgemeinern kann.“<sup>165</sup>

Die Meinungen divergieren bezüglich der Frage, ob man die Bewertung und Überprüfung der Einsätze neuer Medien als Evaluation bezeichnen oder zum selbstverständlichen Bestandteil wissenschaftlichen Arbeitens zählen sollte. Es herrscht Einigkeit darüber, dass Überlegungen und Aktivitäten unternommen werden müssen, um die Qualität der zukünftigen Ergebnisse solcher Projekte zu erhöhen. Vor der Diskussion des Themas *Evaluation von Multimedia* wird der Versuch unternommen, das schwer greifbare Feld *Multimedia* oder *Neue Medien* für das Vorhaben dieser Arbeit einzugrenzen.

Die Evaluation wird im Folgenden aus der allgemeinen und einer speziellen Perspektive betrachtet. Zunächst wird die Evaluation allgemein als Instrument betrachtet, d.h. ohne einen Evaluationsgegenstand zu berücksichtigen. Es werden exemplarisch drei verschiedene Ansätze vorgestellt:

- Scriven unterscheidet verschiedene Formen der Evaluation.
- Beywl stellt das Modell der responsiven Evaluation vor, beschreibt die Weiterentwicklung der Evaluationsmethodologie.
- Abschließend wird das *Vierebenenmodell* von Kirkpatrick zur Evaluation vorgestellt - ein Standardansatz aus den U.S.A., der auch in Deutschland häufig herangezogen wird.

Die selektive Betrachtung dieser drei Ansätze wurde einer umfassenden Diskussion klassischer Evaluationsformen vorgezogen. Die drei allgemeinen Ansätze zur Evaluation werden herangezogen, um einen Raum aufzuspinnen. Durch diesen und die kritische Diskussion der Ansätze wird die Position des Autors in Bezug auf Evaluation herausgestellt. Der allgemeinen Betrachtung folgt die spezielle: die Evaluation von Multimedia bzw. neuer Medien wird diskutiert. Es wird beschrieben, wie sich die Rolle der Evaluation in Zusammenhang mit den neuen Medien verändert, welche neuen Ansprüche gestellt werden.

## 6.1 Evaluation als Instrument - allgemeine Ansätze

**Formen der Evaluation nach Scriven** Im Rahmen der Diskussion um Evaluation von Multimedia bewegen sich viele Wissenschaftler weg von dem Versuch, den Evaluationsbegriff einheitlich zu definieren. Erich Wagner gibt zu bedenken:

„ (...) je schärfer man sich auf Ziele konzentriert, je stärker man sich methodisch einschränkt, um so mehr verliert man Dinge aus dem Blick, die genauso wichtig sein können. D.h. ich schätze nicht so sehr eine relativ verbreitete Auffassung von Evaluation, die besagt, man müsse vorher genau die Ziele des Projektes (...) formulieren und genau daraufhin Antworten suchen, methodisch konsistent und sauber. Es führt oft dazu, dass man Dinge erreicht, geschaffen hat, die sehr wichtig sind, die man aber gar nicht sehen kann, weil man nicht sieht, was man nicht im Blick hat. Und daher finde ich an dieser Stelle ein offeneres Herangehen ganz sinnvoll.“<sup>166</sup>

---

<sup>165</sup>[Kindt1999], S. 128ff.

<sup>166</sup>Vgl. Podiumsdiskussion [Kindt1999].

Der Leser wird in Teil IV erkennen, dass die dort dokumentierte Evaluationsstudie diesem Verständnis von Evaluation folgt. Um sie einordnen zu können, wird der Vorschlag Scrivens<sup>167</sup> betrachtet. Für ihn stellt die eigentliche *Bewertung* das Trennkriterium dar. Er differenziert zwischen den drei Formen. Im Folgenden wird eine Form, die Echte-Evaluation, betrachtet, da sie der Charakterisierung der Evaluationsstudie dieser Arbeit dient.

Als **Echte-Evaluationen** bezeichnet Scriven solche Untersuchungen, die sowohl die Definition, als auch die Beurteilung und Begründung von Wertansprüchen und die darauf basierenden Zieldefinitionen in den Mittelpunkt ihrer Analyse stellen. Echte-Evaluationen sind gekennzeichnet durch Methodenvielfalt sowie interessenausgeglichene Vorgehensweise und Argumentation. Spricht Scriven in diesem Zusammenhang von zielfreier Evaluation, so meint er damit nicht, dass echte Evaluationen in einem interessenneutralen Raum stattfinden. Sie sind durch die klaren Wertansprüche vielmehr interessenorientiert oder interessengeleitet und können daher je nach Orientierung weiter unterteilt werden.

Die von Scriven vorgenommene Differenzierung zwischen verschiedenen Evaluationsformen nach der Qualität ihrer Bewertung bildet die Basis der Evaluationsstudie. Diese erhebt den Anspruch, von Scriven als Echte-Evaluation bezeichnet zu werden. Die allgemeinen Ausführungen Scrivens lassen sich auf die Evaluation von Multimedia übertragen. Vor der speziellen Diskussion der Evaluation von Multimedia wird ein Ansatz aus den U.S.A. zur Evaluation näher betrachtet, der die politische Seite in die Überlegungen zur Evaluation einbezieht.

**Modell der responsiven Evaluation nach Beywl** In einem Artikel aus dem Jahr 1988<sup>168</sup> beschreibt Beywl die Evaluation als eine Disziplin, die ein Bindeglied zwischen Wissenschaft und Politik bilden könnte. Er zeichnet ein umfassendes und detailliertes Bild von der Entwicklung, dem Stand und der Methodologie der Evaluation in den U.S.A. auf. In den Mittelpunkt seiner Ausführungen stellt er das Modell der responsiven Evaluation, die sich in Nordamerika als Evaluationsdisziplin gegen traditionelle Ansätze durchsetzen konnte. Dabei überträgt er die Evaluationsmethodologie aus den U.S.A. nicht undifferenziert auf Deutschland:

„Es ist Bestandteil der Grundüberlegungen der responsiven Evaluation, dass diese eine kulturell gebundene Praxis ist, (...)“<sup>169</sup>.

Statt Erkenntnisse unreflektiert zu übertragen, gibt Beywl neue Anstöße für die Weiterentwicklung der Evaluationsmethodologie in Deutschland.

„Evaluation soll verstanden werden als systematisch vorgehende Disziplin mit originärer Zielsetzung, eigener - sie von Forschung unterscheidener - Logik und damit besonderer Stellung zwischen Wissenschaft und Politik.“<sup>170</sup>

Im Rahmen dieser Arbeit ist es von Interesse, ob es Evaluationsansätze gegeben hat, die unserem Verständnis von Evaluation ähnlich sind, um Erkenntnisse für die Evaluationsstudie<sup>171</sup> ableiten zu können. Ein Ansatz ist die von Beywl aufgestellte *Typologie von Evaluationsmodellen*<sup>172</sup>. Die acht darin enthaltenen Modelltypen der Evaluation sind nicht

<sup>167</sup>Vgl. [Kindt1999], S. 79ff.

<sup>168</sup>Vgl. [Beywl1988].

<sup>169</sup>[Beywl1988], S. 8.

<sup>170</sup>[Beywl1988], S. 13.

<sup>171</sup>Vgl. 14.2.

<sup>172</sup>Vgl. [Beywl1988], S. 45.



en detail von Interesse, sondern nur die durch sie widergespiegelte Herangehensweise an Evaluation. Wir schließen uns Beywls Überlegungen an, dass es bestimmte Typen von Evaluationen gibt, dass diese aus verschiedenen Perspektiven auf zu evaluierende Sachverhalte, Programme ect. schauen und dass sie als Orientierungshilfe dienen können. Ordnet man eine Evaluationsstudie einem dieser Modelltypen zu, kann dies mit Zuordnungsproblemen verbunden sein.

„Modelltypen unterliegen mannigfachen Schwierigkeiten, die auf Unschärfen der jeweiligen Konzeptionen, auf fortlaufende Veränderungen und Überarbeitungen des Modells ... beruhen.“<sup>173</sup>

Im Rahmen dieser Arbeit stehen Evaluations- bzw. Einsatzkonzepte im Mittelpunkt der Betrachtungen. Auch Beywl vertritt die Ansicht, dass Evaluationen fest konzipiert und gleichzeitig modifizierbar sein müssen.

„Der Entwurf einer Evaluationsstudie ist eine Kunst. Bei jeder weiteren Unternehmung muss das Design aufs neue gewählt werden, und die Wahlmöglichkeiten sind beinahe unzählig.“<sup>174</sup>

Seine Ausführungen machen deutlich, dass man Modelle benötigt, um von der Vielzahl existierender Evaluationen zu abstrahieren. Diese werden trotz angedeuteter Unschärfen in der Abgrenzung in einer Typologie von Evaluationsmodellen zusammengefasst. Diese Zusammenfassung ähnlicher Evaluationsmodelle wird benötigt, um zwischen Evaluationsmodellen differenzieren zu können. Allerdings darf man nicht vergessen, dass es sich um Modelle handelt. Modelle können Ansätze, Gestaltungs- oder Strukturvorschläge, Hauptfragestellungen, Arbeitsschritte, Instrumente/Methoden, angestrebte Ergebnisse und Ziele enthalten.

Ein weiterführender Ansatz Beywls besteht in der kritischen Betrachtung sogenannter *Standards für Evaluationen*<sup>175</sup>. Schon Mitte der sechziger Jahre gab es die ersten Versuche, Evaluationen anhand von Qualitätsrichtlinien zu beurteilen. Diese konnten sich allerdings nicht durchsetzen. Anfang der achtziger Jahre wurden in den U.S.A. verschiedene Standardkataloge für Evaluationen veröffentlicht. Die Standards des Joint-Committees<sup>176</sup> werden im Folgenden näher betrachtet.<sup>177</sup> Das Committee präsentiert 30 Standards. Diese werden in die folgenden vier Gruppen eingeteilt:

1. **Nützlichkeits-Standards** sollen sicherstellen, dass die Evaluation sowohl die Interessen der Beteiligten als auch die der Verantwortlichen berücksichtigt. Außerdem soll das Zustandekommen der Evaluationsergebnisse nachvollziehbar sein, zugrunde liegende Werturteile offengelegt werden.
2. **Anwendbarkeits-Standards** sollen eine kosteneffektive Evaluation gewährleisten und dafür sorgen, dass diese realistisch geplant wird. Durch Planung und Anwendung praktikabler Verfahren sollen größere Unterbrechungen des Evaluationsprozesses vermieden werden.

---

<sup>173</sup>[Beywl1988], S. 44.

<sup>174</sup>[Beywl1988], S. 44.

<sup>175</sup>Vgl. [Beywl1988], S. 113ff.

<sup>176</sup>Das "Joint Committee on Standards of Educational Evaluation" wurde 1976 ins Leben gerufen, arbeitete 5 Jahre an diesen Standards, ist kein Berufszusammenschluss aus Evaluatoren, sondern versteht sich auch als Vertreter der Interessen möglicher Nutzer. Die Standards werden von 14 amerikanischen Mitgliedsorganisationen getragen.

<sup>177</sup>Vgl. [Beywl1988], Übersicht auf S. 118.

3. **Korrektheits-Standards** zielen darauf ab, Evaluationsethik zu verankern. Wichtige Evaluationsaktivitäten sollen schriftlich zwischen allen Beteiligten festgelegt werden. Diese sollen als Basis bei Konflikten herangezogen werden.
4. **Genauigkeits-Standards** sollen erreichen, dass besonders technische Informationen über den zu evaluierenden Gegenstand (Programme, Materialien) jederzeit verfügbar sind. Außerdem sollen der Gegenstand, die Ziele sowie Informationsquellen deutlich identifiziert und beschrieben werden. Es soll während des gesamten Evaluationsprozesses auf valide und reliable Messung, begründete Schlussfolgerungen und unverzerrte Berichterstattung geachtet werden.

Die Betrachtung einiger Stellungnahmen zu den 30 Standards hat ergeben, dass diese zwar grundsätzlich positiv bewertet wurden, aber folgende Lücken enthalten: Es ist unklar, wie man sich bei Konflikten von zwei oder mehreren Standards verhalten soll. Es fehlt ein Standard zur Bedürfnisbestimmung sowie Material, mit dessen Hilfe diese standardorientierte Evaluation vermittelt und erlernt werden kann.

Die Standards wurden bei der Konzeption der in Teil IV vorgestellten Evaluationsstudie berücksichtigt. Sie wurden allerdings auf die spezifischen Besonderheiten und Schwerpunktsetzungen der Evaluationsstudie abgestimmt. Die Arbeit richtet sich somit nicht generell gegen Standards für Evaluationen, betrachtet diese jedoch nur als Ausgangsbasis. Ein Problem besteht in der Messung der Erreichung der Standards. Im Folgenden wird ein weiterer Ansatz beschrieben, zwischen verschiedenen Evaluationen zu differenzieren.

**Das Vierebenenmodell von Kirkpatrick** Dem Leser wird ein vierstufiges Modell der Evaluation von Trainingsprogrammen vorgestellt, das *Vierebenenmodell* von Kirkpatrick aus dem Jahr 1994. Es zählt heute noch zu den Standardansätzen von Evaluationen in den U.S.A. und wird auch in Deutschland häufig herangezogen.<sup>178</sup>

Die folgenden drei Ebenen dieses Modells sind für die Evaluationsstudie dieser Arbeit relevant:

1. **Reaction:** Die Evaluation konzentriert sich auf den Teilnehmer. Kirkpatrick stellt die These auf, dass eine positive Reaktion eine gute Voraussetzung für erfolgreiches Lernen darstellt, eine negative dieses erschwert. Es werden keine Thesen der Art „je besser die Reaktion, desto höherer Lernerfolg.“ formuliert.
2. **Learning:** Das Ziel dieser Ebene besteht in der Evaluation des Lernens der Teilnehmer. Lernen beschreibt Kirkpatrick als die Erweiterung von Kenntnissen, die Erhöhung von Fertigkeiten, die Veränderung von Einstellungen. Diese Ebene kann durch die Konstruktion von Wissen im Sinne des Konstruktivismus ergänzt werden. Der Studierende lernt, indem er vorhandenes und neues Wissen kombiniert, um Probleme zu lösen.
3. **Behavior:** Auf dieser Ebene steht das Verhalten der Teilnehmer im Vordergrund. Es soll beobachtet werden, wie sich ihr Verhalten in Reaktion auf das neu erworbene Wissen verändert. Es wird untersucht, inwiefern die Veränderung des Wissens eine Veränderung des Verhaltens bewirkt.

Betrachtet man diesen Ansatz als Basis, so muss ein gewisses Bildungsangebot bereits vorhanden sein. Dieses wird hinsichtlich seiner Wirkungen untersucht. Kirkpatrick lagert die Entwicklungsphase also indirekt vollständig vor diesen vier Ebenen an.

---

<sup>178</sup>Vgl. Schenkel [Schenkel2000], S.59ff.

## 6.2 Evaluation von Multimedia - spezielle Ansätze

Im Folgenden wird herausgestellt, dass dem Evaluationsbegriff im Zusammenhang mit den neuen Medien eine neue Rolle und Bedeutung zukommt. Eine kritische Auseinandersetzung mit existierenden Evaluationen von Multimedia verdeutlicht den eigenen Standpunkt auf diesem Gebiet. Erich Wagner fasste im Rahmen der Evaluation von Multimedia entscheidende Fragen zusammen:

„Evaluation kann man inszenieren, weil man etwas lernen möchte. Dann sind für mich die entscheidenden Fragen: Wie fange ich das an? Wie systematisch mache ich es? Wie kontinuierlich mache ich das? Welche Formen finde ich? Welche Regeln beachte ich dabei? Welche Fehler vermeide ich nach Möglichkeit?“<sup>179</sup>

Es wird der Versuch unternommen, auf diese Fragen Antworten zu geben. In diesem Zusammenhang ist es notwendig, zu definieren, was unter der Evaluation von Multimedia zu verstehen ist. Es wird im Kern der Frage nachgegangen werden, welche Aspekte und Komponenten bei der Suche nach erfolgreichem Lernen mit multimedialer statistischer Lernsoftware eine Rolle spielen.

**Phasenweise Evaluation** Im klassischen Sinn stand die Evaluation immer am Ende eines Entwicklungsprozesses. Bei der Evaluation von Multimedia hingegen, d.h. der Bewertung und Überprüfung multimedialer Lernprogramme, hat sich folgende Erkenntnis durchgesetzt: Der Entwicklungs-, Einsatz- und Evaluationsprozess sollte in verschiedene Phasen unterteilt werden. Nach Abschluss jeder Phase sollte eine Bewertung erfolgen.

„For the educational technology field, evaluation was now being viewed as an integral and ongoing part of the instructional development process.“<sup>180</sup>

Dazu ist es notwendig, den Entwicklungsprozess multimedialer Lernsoftware in Phasen zu unterteilen und für jede Phase bestimmte Ziele zu definieren. Im Rahmen der Evaluation der jeweiligen Phase besteht die Möglichkeit, die Zielerreichung zu überprüfen.

„Qualitätskontrolle findet nach neuerem Verständnis nicht nur bei der Bewertung des Endproduktes, sondern schon auf allen Stufen des Entwicklungsprozesses statt.“<sup>181</sup>

Werden die Ziele einer Phase nicht erreicht, kann innerhalb der Evaluationsstudie reagiert werden. In der Evaluationsstudie dieser Arbeit wurde phasenweise evaluiert. Das Vorgehensmodell, das für die Entwicklung der Lernsoftware *Statistik interaktiv komplett* und des Einsatzkonzeptes aufgestellt wurde, sieht Rückkopplungen der Phasen vor und ermöglicht so einen iterativen Entwicklungsprozess. Die parallele Betrachtung von Entwicklung und Einsatz ist auch in der folgenden Position Tergans vorhanden.<sup>182</sup> Er unterscheidet bezüglich Multimediasoftware zwischen drei Phasen: der Planungsphase, der Entwicklungsphase und der Einsatzphase. Das Evaluationsziel der Entwicklungsphase beschreibt er wie folgt:

„Ziel der Evaluation ist die Sicherung und Optimierung der pädagogischen Qualität des jeweiligen Bildungsangebots unter Berücksichtigung der gewählten

<sup>179</sup>[Kindt1999], S. 130.

<sup>180</sup>[RossMorrison2000].

<sup>181</sup>[Kindt1999], S. 109.

<sup>182</sup>Vgl. [Schenkel2000], S. 22-51.

Ziele, Qualitätskriterien und Rahmenbedingungen des vorgesehenen Einsatzes. Die Qualität des sich in der Entwicklung befindlichen Bildungsangebotes wird in dieser Phase wiederholt überprüft, damit eventuelle Mängel behoben werden können.“<sup>183</sup>

Das Evaluationsziel der Einsatzphase definiert er wie folgt:

„Ziele der Evaluation in der Einsatzphase sind die Erfassung der Wirkungen eines Bildungsangebotes (Akzeptanz, Lernerfolg, Wissenstransfer), seiner Effizienz verglichen mit alternativen Maßnahmen sowie das Kosten-Nutzen-Verhältnis, d.h. der Aufwand, mit dem bestimmte Veränderungen erreicht wurden.“<sup>184</sup>

Tergan berücksichtigt zudem den Aspekt, dass die Evaluation in der Einsatzphase auch dann durchgeführt werden kann, wenn das Multimediaprodukt noch nicht fertiggestellt ist. „Nicht selten werden Bildungsangebote erst unter Erstbedingungen erprobt und schließlich auch weiterentwickelt.“<sup>185</sup> Dies galt für den Einsatz und die Evaluation von *Statistik interaktiv komplett*: Erkenntnisse der Evaluation in der Einsatzphase wurden auf die Entwicklungs- oder Weiterentwicklungsphase übertragen, um das Produkt entsprechend der Nutzerbedürfnisse zu gestalten und eine qualitativ hochwertige Lernsoftware zu entwickeln.

In der Evaluationsstudie dieser Arbeit wurde phasenweise evaluiert. Dies betraf die Entwicklung der Lernmodule und des Einsatzkonzeptes. Dem Leser wird in Teil III der vorliegenden Arbeit der phasenweise Entwicklungsprozess anhand der Komponente *Statistiklabor* demonstriert. Zudem werden die Schwerpunkte der Phasen vorgestellt.

**Theoretischer Definitionsversuch** Das Ziel besteht in der Herausarbeitung eines theoretischen Definitionsversuches zur Evaluation von Multimedia. Baumgartner unternimmt in seinem Artikel *Evaluation mediengestützten Lernens, Theorie-Logik-Modelle*<sup>186</sup> den Versuch einer umfassenden theoretischen Definition zur Evaluation mediengestützten Lernens. Er charakterisiert verschiedene Felder expliziter Definitionsversuche. Anhand der Beschreibung und Bewertung der einzelnen Felder wird die dieser Arbeit zugrunde liegende Position zur Evaluation von Multimedia herausgearbeitet.<sup>187</sup>

1. **Die relativistische Position:** Evaluation lässt sich nicht definieren. Baumgartner fasst darunter solche Ansätze, in denen gar nicht erst der Versuch einer Definition unternommen wird. Vertritt man die Meinung, Evaluation ließe sich nicht definieren, unterbindet man damit die Beantwortung wissenschaftlicher Forschungsfragen nach der Weiterentwicklung der Evaluationspraxis und der wissenschaftstheoretischen Bewertung verschiedener Evaluationsmodelle. Die kritiklose Auflistung und Aneinanderreihung verschiedener Definitionsversuche zur Stützung der These, dass Evaluation zu komplex und vielschichtig ist, um sinnvoll abgegrenzt zu werden, kann der Meinung des Autors nach nicht als wissenschaftliches Arbeiten verstanden werden.

---

<sup>183</sup>[Schenkel2000], S. 40.

<sup>184</sup>[Schenkel2000], S. 42.

<sup>185</sup>[Schenkel2000], S. 42.

<sup>186</sup> erschienen in [Kindt1999].

<sup>187</sup>Die folgenden Ausführungen enthalten die jeweiligen vom Autor ausgewählten wichtigsten Aspekte der einzelnen Felder, an der die in dieser Arbeit vertretene Position zur Evaluation durch positive/negative Kritik verdeutlicht wird.

2. **Evaluation als quantitatives Analyseverfahren:** Hier wird Evaluation implizit gleichgesetzt mit der Konstruktion und Auswertung von Tests. Es werden fast ausschließlich quantitative Messverfahren betrachtet und ihre statistischen Probleme beschrieben.  
Die Kritik liegt in dem völligen Fehlen eines qualitativen Ansatzes. Bei Evaluation kann es nicht bloß um eine systematische Datenerfassung und ihre quantitative Auswertung gehen. In dieser Arbeit wird die Notwendigkeit der Kombination qualitativer und quantitativer Evaluationsverfahren zur Bewertung multimedialer Lernsoftware gesehen. Daher wird dieser Ansatz als nicht ausreichend eingestuft.
3. **Evaluation als eine Anwendung von Methoden:** Vertreter dieses Evaluationsansatzes setzen sozialwissenschaftliche Methoden um, beschreiben und diskutieren sie, verwenden auch qualitative Ansätze. Die Evaluation nur als ein Bündel von Methoden aufzufassen, reicht Baumgartner nicht, er ist nicht der Auffassung, dass Evaluation mit Methodenlehre gleichzusetzen ist, die auf eine bloße objektive Beschreibung von Sachverhalten bzw. Evaluanten hinausläuft. Es muss bemerkt werden, dass der Ansatz noch nicht ausreichend ist, da die eigentliche Bewertung vernachlässigt wird.
4. **Evaluation als Bewertung:** Dieses Feld umfasst das, was in dieser Arbeit unter Evaluation verstanden wird. Der Evaluationsbegriff wird durch den Prozess der Bewertung charakterisiert.

"Evaluation is the process of determining the merit, worth and value of things, and evaluations are the products of that process."<sup>188</sup>

Dabei wird nicht festgelegt, welche Evaluationsverfahren zum Einsatz kommen. Die Kombination qualitativer und quantitativer Evaluationsverfahren ist somit möglich. Die Charakterisierung der Evaluation als Prozess kommt dem der phasenweisen Evaluation gleich. Die aufeinander folgenden Phasen determinieren den Evaluationsprozess.

Baumgartner hat sich bei seinem Definitionsversuch auf die Evaluation an sich bezogen. Die Ausführungen haben gezeigt, dass dieser auf die Evaluation multimedialer Lernsoftware übertragbar ist.

Die bisherigen Ausführungen stellen folgende Anforderungen an die Evaluation. Sie muss Scrivens Definition der *echten* Evaluation, den Standards des Joint Committees, sowie den drei diskutierten Ebenen des Modells von Kirkpatrick Genüge leisten. Baumgartners Charakterisierung der Evaluation als Bewertung ergänzt diese Ansätze. Die Position dieser Arbeit wird durch den aufgespannten Raum dieser Ansätze eingegrenzt.

**Evaluation der Qualität multimedialer Lernsoftware** Die Evaluation von Multimedia ist in die Diskussion gekommen, da sich viele nichts sagende Abschlussberichte bei den Projektträgern sammeln. Ein häufig anzufindener Schlusssatz lautet: *Wie das Projekt gezeigt hat, ist weitere Forschung erforderlich.* Aus dieser Situation heraus ist die Verpflichtung der Projektnehmer zur Selbstevaluation erwachsen. Erich Wagner hält die Selbstevaluation für notwendig, ruft seine Kollegen dazu auf, sie sinnvoll durchzuführen, sich auf grundlegende Regeln zu einigen. Keil-Slawik sagte in einer Podiumsdiskussion:

---

<sup>188</sup>[Scriven1996].

„ (...) weil nämlich in der Vergangenheit genauso viel Schindluder mit schlechter Qualität oder schlechter Multimedia wie mit schlechter Evaluation getrieben worden ist. Wenn ich heute höre, dass aus einem Projekt behauptet wird, wir haben die Lernleistung um 30 Prozent gesteigert, halte ich das für Quatsch.“<sup>189</sup>

Allerdings befindet sich die Evaluation neuer Medien in einem kontinuierlichen Lernprozess. Die Entwicklung dieses Lernprozesses stuft der Autor positiv ein: Es wurde erkannt, dass traditionelle Evaluationsansätze an Grenzen stoßen und deshalb evolutionäre Gestaltungsansätze erforderlich sind. Evolutionäre Gestaltungsansätze bestehen aus aufeinander folgenden Zyklen von Entwicklung, Einsatz, Bewertung und Weiterentwicklung. Es wurde eingesehen, dass eine abstrakte inhaltliche Analyse von Software nicht aussagekräftig ist. Diese kann die Inhalte der Lernsoftware zwar mit der klassischen Lehrbücher vergleichen, Aussagen über die Wirkung der Komponenten oder die Kontinuität des Lernprozesses sind allerdings nicht möglich.

### Evolutionäre Gestaltungsansätze

Die evolutionären Gestaltungsansätze zeigen nicht die Überlegenheit *eines* favorisierten Ansatzes. Sie versuchen, vielseitige Lösungsansätze zu entwickeln und zu etablieren, die das übergeordnete Ziel verfolgen, Antworten auf die entscheidenden Fragen zu geben:

- Wie kann durch eine lernförderliche Infrastruktur der Aufwand für die Lehrenden reduziert werden?
- Wie können Projekte so evaluiert werden, dass Erfahrungen und Ergebnisse einen Qualitätsnachweis und vergleichbare Anhaltspunkte für eine Verallgemeinerung liefern?
- Wie können sich durch neue Medien neue Qualitäten ergeben?

Die neuen Ansätze konzentrieren sich auf mehrere Felder und verbinden Evaluation mit dem Einsatz der neuen Medien und darauf abgestimmte Entwicklungsprozesse. Sie bewerten die Qualität multimedialer Lernsoftware nach Kriterien der folgenden Art: Modifizierbarkeit, Verknüpfbarkeit, Wiederverwendbarkeit und durchgängige Verfügbarkeit von Materialien unterschiedlicher Art.

Die entwickelte Lernsoftware *Statistik interaktiv komplett* wurde im Hinblick auf diese Kriterien entwickelt. Sie kann als ein Steinbruch für das Statistiklernen angesehen werden. Dies gilt insbesondere für die Komponente *Statistiklabor*. Dort werden verschiedene Instrumente bereitgestellt, um statistisch zu experimentieren. Der Studierende kann das Material oder die Objekte nutzen, die er zu seiner Wissenskonstruktion im Sinne des Konstruktivismus benötigt. Der Einsatz der einzelnen Komponenten und Module der Lernsoftware, d.h. das entwickelte Einsatzkonzept, stellt den Aspekt des Steinbruchs in den Vordergrund. Dem Studierenden werden verschiedene Komponenten angeboten, die miteinander verknüpft eingesetzt werden können. Die Statistiklaborseiten sind modifizierbar und wiederverwendbar gestaltet. Ein Nachteil von *Statistik interaktiv komplett* ist die fehlende Modifizierbarkeit der anderen Komponenten. Die Theoriekomponenten können vom Dozenten nicht durch eigene Materialien ergänzt werden. Gleiches gilt für die Beispiel- und Aufgabenkomponenten. Diese Nachteile haben sich ergeben, obwohl die evolutionären Gestaltungsansätze berücksichtigt wurden.

---

<sup>189</sup>[Kindt1999], S. 132.

Im Laufe der letzten Jahre wurden Einsichten und Erfahrungen in Bezug auf den alltags-tauglichen Einsatz von Multimedia bzw. neuen Medien gesammelt, die einen Wechsel in der Entwicklungsperspektive zur Folge hatten. Es wird nicht mehr nach dem "One-Best-Way" in der Entwicklung und Systemgestaltung gesucht, sondern nach Entwicklungsstrategien, bei denen viele verschiedene Entwicklungsparameter und Rahmenbedingungen eine Rolle spielen. Das heißt, dass sich zuverlässige Entwicklungsstrategien erst im Laufe der Zeit herausbilden können, es sich bei den Ansätzen der evolutionären Systemgestaltung um Lernprozesse handelt, die noch nicht abgeschlossen sind, dass der Einsatz und die Evaluation von Multimedia selbst der Lernprozess ist. Man befindet sich im deutlichen Gegensatz zu der klassischen Evaluation, die „das politische Moment der Rechtfertigung der eingesetzten Mittel und damit der Bestätigung, den Stein der Weisen oder zumindest relevante Splitter schon gefunden zu haben oder aber zumindest über vergleichbare Qualitätsmaßstäbe zu verfügen, enthält.“<sup>190</sup> Dieser Lernprozess soll durch den Einsatz und die Evaluation von *Statistik interaktiv komplett* positiv beeinflusst werden.

**Evaluationskriterien und Kriterienkataloge** Im Rahmen der Qualitätsüberprüfung neuer Medien gibt es Vertreter, die Kriterienkataloge entwickeln, um multimediale Lernsoftware einheitlich zu bewerten. Die Bewertung multimedialer Lernsoftware bezüglich bestimmter festgelegter, objektiver Evaluationskriterien muss kritisch gesehen werden. Die Kritik fängt damit an, dass Variablen ausgewählt werden, anhand derer eine Überprüfung bestimmter Kriterien unmöglich ist.

Issing gibt ein Beispiel, das diesen Aspekt verdeutlicht:

„Wenn beispielsweise bestimmte Fähigkeiten zur Beherrschung eines Textverarbeitungssystems erlernt werden sollen, dann reicht es nicht aus, die Akzeptanz des Programms zu messen und die Kenntnis der Funktionstasten abzutprüfen. Notwendig wird in diesem Fall eine genaue Analyse des Lernverhaltens bei neuen Aufgaben.“<sup>191</sup>

Es wurde bereits der Versuch unternommen, Checklisten zu entwerfen, anhand derer Lernsoftware bezüglich des Vorhandenseins bestimmter lernwirksamer Parameter überprüft werden sollte. Von empirisch validierten Checklisten ist man allerdings noch weit entfernt. Die Schwierigkeit besteht grundsätzlich darin, Evaluationskriterien festzulegen, die kontextunabhängig geeignet sind, multimediale Lernsoftware zu beurteilen. Man muss sich die Frage stellen, ob dies überhaupt annähernd möglich sein kann. Stellt man die These auf, dass die Qualität einer Lernsoftware zu großen Teilen davon abhängt, wie aktiv der Studierende mit ihr lernt, so wird diese Qualität erst durch die konkrete Anwendungssituation hergestellt. Versteht man unter der Qualität einer multimedialen Lernsoftware z.B. die mit ihr erreichbaren Lernerfolge oder erreichbaren Handlungskompetenzen, also Größen, die erst im Prozess selbst konstruiert werden, so erscheinen objektive Evaluationskriterien ungeeignet. Der entscheidende Aspekt, der in jedem Versuch, multimediale Lernsoftware zu evaluieren, berücksichtigt werden muss, ist der, dass subjektive Lernerfolge der Studierenden (hergestellt durch eigenes Lernen und durch Lehrende gesteuertes Lernen mit der Lernsoftware) in keinem sinnvollen Zusammenhang zu objektiven Kriterien stehen können.

---

<sup>190</sup>[Kindt1999], S. 11.

<sup>191</sup>[IssingKlimsma1997].

Betrachten wir weitere Standpunkte in der Diskussion um Evaluationskriterien - Schenkel spricht von der Qualitätsbeurteilung durch Kriterienkataloge:

„Der Einsatz von Kriterienkatalogen bei der Evaluation von Lernsoftware ist häufig mit der Erwartung verbunden, dass es einige wenige kritische Erfolgsfaktoren gebe, die eine überzeugende Einschätzung von Lernprogrammen ermöglichen.“<sup>192</sup>

Dies bezweifelt Schenkel und spricht sich gegen eine Evaluation auf Basis von Kriterienkatalogen aus. In der durchgeführten Evaluationsstudie wird als Gegensatz zu Kriterienkatalogen unter anderem auf Basis von Befragungen der Studierenden evaluiert. Tergan<sup>193</sup> stellt heraus, dass Befragungen darauf abzielen, Informationen bezüglich der Akzeptanz, des Lehrangebots, der pädagogischen Qualität und didaktischen Gestaltung zu erheben. Differenziert man zwischen offenen und geschlossenen Fragen, so kommen die geschlossenen Fragen den sehr kontrovers diskutierten Kriterienkatalogen oder Checklisten sehr nahe. Offene Fragen sind dadurch gekennzeichnet, dass sie dem Befragten eine freie Wahl bezüglich der Form, des Inhaltes und der Ausführlichkeit seiner Antwort lassen. Geschlossene Fragen hingegen schränken diese Möglichkeiten durch die Art ihrer Formulierung stark ein. In der Evaluationsstudie wurden offene und geschlossene Fragen kombiniert. Die offenen Fragen führten zu besseren Ergebnissen, ihre Auswertung war allerdings mit viel Aufwand verbunden.

Fricke stellt die These auf, dass die Evaluation auf Basis von Kriterienkatalogen deshalb bei vielen Projekten beliebt ist, weil sie aufwendige Evaluationsstudien ersetzt. Er stellt heraus, dass ein Großteil von Katalogen Kriterien enthält, die empirisch nicht überprüft sind und überwiegend technische Aspekte behandeln. Wünschenswert und qualitativ hochwertiger wären seiner Meinung nach didaktische Kriterien, die aus Lehr- und Lerntheorien abgeleitet wären. Die herkömmlichen Kriterienkataloge haben ihre praktische Berechtigung nur in drei Punkten:<sup>194</sup>

- Als k.o.-Kriterien können sie für eine schnelle und kostengünstige Vorselektion von Lernprogrammen benutzt werden.
- Sie können dazu dienen, dem praktischen Anwender die Komplexität der Wirkungsfaktoren multimedialer Lernumgebungen zu verdeutlichen.
- Sie helfen Praktikern, eigene Hypothesen über Lernen mit Multimedia aufzustellen, geben positive Anregungen für Evaluationsstudien und Einsätze von Multimedia.

Betrachtet man die existierenden Kriterienkataloge aus dieser Perspektive, so sind sie in einem gewissen Rahmen und unter diesen Vorbehalten anwendbar. Fricke stellt Gründe vor, die gegen eine Evaluation auf Basis von Kriterienkatalogen sprechen. Die folgenden zwei sind für die Evaluationsstudie von Bedeutung:<sup>195</sup>

*Geringe praktische Signifikanz der Qualitätskriterien:* Sucht man lernwirksame Kriterien, so muss beachtet werden, dass der Lernerfolg des multimedialen Lernprogrammes von vielen Kriterien und auch den Wechselwirkungen zwischen diesen abhängig ist. Ein einziges Kriterium allein wird den Gesamtlernerfolg nur sehr gering beeinflussen. Die Höhe der

---

<sup>192</sup>[Schenkel2000], S. 63.

<sup>193</sup>Vgl. [Schenkel2000], S. 30ff.

<sup>194</sup>Vgl. [Schenkel2000], S. 87.

<sup>195</sup>Vgl. [Schenkel2000], S. 76ff.



Korrelation zwischen den Merkmalen Lernerfolg und Ausprägung des Kriteriums wird sehr gering sein. Die Anwendbarkeit von Kriterienkatalogen, bei denen ein Kriterium nach dem anderen abgehandelt wird, ist somit fragwürdig. An dieser Stelle kann man erkennen, wie wichtig der Bezug zu etablierten Lehr- und Lerntheorien ist. Auf der Suche nach aussagekräftigen, validen Qualitätsmerkmalen und somit Kriterien bieten diese eine wissenschaftlich bewiesene Grundlage. Merkmale werden dort bereits hinsichtlich ihrer Beeinflussung des Lernerfolges untersucht. Dadurch kann vermieden werden, eine unstrukturierte Anhäufung nicht valider und somit unbrauchbarer Kriterien vorzunehmen.

*Nichtberücksichtigung des Verwertungszusammenhanges einer Bildungssoftware:* Soll eine multimediale Lernsoftware zielführend evaluiert werden, so muss dies innerhalb eines bestimmten Einsatzszenarios geschehen. Stellt man diese These auf und folgen ihr, so ist es unwahrscheinlich, dass allgemeingültige Kriterienkataloge spezielle Besonderheiten in bestimmten Szenarien erfassen können. Die Qualitätsmerkmale werden innerhalb der einzelnen Szenarien variieren, es werden unterschiedliche Wechselwirkungen entstehen, das Lernergebnis wird in verschiedener Art beeinflusst werden. Die Evaluationsstudie untersucht vorwiegend Wechselwirkungen verschiedener Komponenten. Eine Bewertung auf Basis von Kriterien ist daher problematisch.

Aus der Diskussion geht hervor, dass es gerade im Bereich der neuen Medien keine Kriterienkataloge gibt, die eine ausreichende Basis in der theoretischen und empirischen Forschung besitzen. Hier wird die These vertreten, dass Evaluationskriterien nicht kontextunabhängig aufgestellt werden können. Der Kontext der durchgeführten Evaluationsstudie ist die Statistik. In Teil IV dieser Arbeit werden multimediale und traditionelle Komponenten hinsichtlich ihrer Eignung überprüft, konkreten Defiziten der Statistikausbildung entgegenzuwirken. Das Ziel dieser Arbeit besteht nicht darin, starre Kriterienkataloge für die Evaluation neuer Medien zu entwickeln, sondern individuelle Kriterien für konzipierte Einsätze abzuleiten und empirisch hinsichtlich ihrer Eignung zu bewerten. Ein Evaluationskriterium ist die Interaktivität. Dies wurde aus den identifizierten Defiziten und den sich daraus ergebenden Zieleigenschaften der Studierenden abgeleitet. Nur im Hinblick auf diese Zieleigenschaften kommt der Interaktivität eine zentrale Rolle zu. Dies ist nicht bei jeder Evaluationsstudie zu multimedialer Lernsoftware der Fall.

**Interaktivität als Evaluationskriterium** Betrachtet man zunächst einige Zitate zur Interaktivität, die ihren Stellenwert in der derzeitigen Diskussion widerspiegeln:

„Man evaluiert dann nicht mehr ein Medium, sondern ein System, das bestimmte Verarbeitungsprozesse beim Lerner in Gang setzen soll.“<sup>196</sup>

„Was Multimedia von allen anderen Medien unterscheidet, ist offenbar der hohe Grad der Interaktion, (...) Borsook (1991) argumentiert, dass das beste Unterscheidungsmerkmal für Multimedia ihr Potential für Interaktivität sei.“<sup>197</sup>

„We must conclude that the point is not: interaction yes or no. The point is: more or less.“<sup>198</sup>

---

<sup>196</sup>[Weidenmann1986].

<sup>197</sup>[Schulmeister2002] S.41.

<sup>198</sup>[Schulmeister2002] S.44.

„Der Begriff „Interaktivität“ ist sicher einer der im Zusammenhang mit modernen Medien sowohl am häufigsten gebrauchten, als auch einer der am wenigsten scharf definierten Begriffe (...) Unter anderem dadurch bedingt wird dieser Begriff in geradezu inflationärer Weise als Beschreibungsmerkmal für fast jede auf dem Markt befindliche Software benutzt und z.T. missbraucht.“<sup>199</sup>

Die Interaktivität ist eine zentrale neue Eigenschaft der neuen Medien. Sie eröffnet neue Möglichkeiten für die individuellen Lernprozesse. Zum Beispiel entstehen neue Kommunikationswege, Informationen oder Materialien können innerhalb eines multimedialen Lernsystems bereitgestellt werden, sofern dies zielgerichtet gestaltet ist. Dies eröffnet die Möglichkeit des direkten Feedbacks beider Seiten. Der Lernende kann Informationen abrufen, dem Dozenten mitteilen, wie er mit diesen zurechtkommt. Der Dozent kann gelöste Übungen, Tests (hier: Übungsaufgaben und statistische Reports) zur Korrektur entgegennehmen, der Studierende braucht diese nur auf einer vorher vereinbarten Kommunikationsplattform zu hinterlegen.

Riehm/Wingert unterscheiden zwischen verschiedenen Interaktionsformen:<sup>200</sup>

1. „Interaktion als die unumgängliche Bedienung eines Programms, vom Aufrufen des Programms selbst bis zu dessen einzelnen Funktionen wie Blättern, Suchen, (...)“
2. Interaktion als die Selektion aus einem u.U. umfänglichen Angebot und insoweit Mitwirkung, Inhalt und Form selbst. Es gibt - denken wir an einen vielfach verzweigten Hypertext - zwischen Inhalt und Form keine fixierte, prä-koordinierte Beziehung mehr; Inhalt ist, was sich der Benutzer bzw. Leser auf der Leseroute zusammenliest.
3. Interaktion schließlich als das mehr oder weniger zielgerichtete Einsetzen von Verarbeitungsroutinen, die z.B. aus Daten eine Grafik erstellen, aus einem eingegebenen Wortgut ein Gedicht machen oder aus einem Text automatisch ein Abstract erzeugen.“

Die zweite Definition kommt dem Verständnis von Interaktion dieser Arbeit nahe. Es setzt implizit die konstruktivistischen Prinzipien Wissenskonstruktion und Lernen durch eigene Aktivität voraus.

Stellt man die Interaktionsmöglichkeiten eines traditionellen Buches denen einer multimedialen Lernsoftware gegenüber, so wird deutlich, dass die Zugriffs-, Analyse- und Manipulationsmöglichkeiten enorm ausgeweitet wurden. Dies führt aber nicht automatisch dazu, dass die Inhalte adäquat gelernt und rezipiert werden. In dieser Arbeit wird die These aufgestellt, dass Interaktion nur dann positiv wirken kann, wenn das aktive Tun des Lernenden im Sinne des Konstruktivismus gefördert und erhöht wird. Dies wird im Einsatzkonzept für die Lernsoftware *Statistik interaktiv komplett* berücksichtigt. Interaktivität wird als konstruktivistisches Prinzip betrachtet, das den Lernprozess positiv beeinflussen kann. Rückt man den Zusammenhang zwischen Lernprozess und Interaktivität in das Zentrum, so ergeben sich daraus unmittelbar Konsequenzen für die Entwicklung und Evaluation. Bei *Statistik interaktiv komplett* wurden im Entwicklungsprozess zahlreiche Interaktionsmöglichkeiten geschaffen. Diese betrafen insbesondere das Statistiklabor.<sup>201</sup>

---

<sup>199</sup>[Pausch1993].

<sup>200</sup>[RiehmWingert1995], S. 201.

<sup>201</sup>Vgl. 10.

Im Rahmen der Diskussion um Evaluation von Interaktivität wird nicht selten die These aufgestellt, dass die Evaluation interaktiver Medien umfangreicher ist als die traditionelle Evaluation. Betrachten wir daher abschließend einige Besonderheiten der neuen Medien und der durch sie angestoßenen Lernprozesse, die zu einer höheren Komplexität des Evaluationsprozesses führen können:

Es entsteht ein neuer individueller Zugang zum Lernen. Die Evaluation muss Wege finden, das nicht mehr direkt beobachtbare Lernen mit der interaktiven Lernsoftware zu evaluieren.

Die Lernenden haben zudem eine heterogene Ausgangssituation und Bedürfnisstruktur. Sie können Lernprozesse individuell gestalten, da die neuen Medien sie weniger Restriktionen unterwerfen. Die neuen Gegebenheiten führen zu unterschiedlichen Zielvorstellungen, Bedürfnis- und Motivationsstrukturen. Es entstehen individuelle Lernstrategien und -ziele und damit das Problem der Messung des individuellen Lernerfolgs/Lernprozesses.

Die Evaluation muss sowohl die statische Qualität der mediengestützten Materialien beurteilen als auch deren Eignung, neue Lernprozesse in Gang zu setzen. Des Weiteren muss die Evaluation berücksichtigen, dass eine Multimediasoftware die individuelle Zusammenstellung der Inhalte ermöglicht, auch die Ebene des Lernen organisiert. Bezüglich der Lernorganisation muss bei der Evaluation von Interaktivität unterschieden werden zwischen den Effekten der Verteilung des Materials und den dynamischen Inhalten bzw. der interaktiven Didaktik des Materials. Die Evaluation muss auf die Trennung dieser Effekte und der sie bestimmenden Faktoren achten, da die neuen Medien beide hervorrufen. Die neuen didaktischen Möglichkeiten wie auch die neuen dynamischen Verteilungsmöglichkeiten müssen gleichermaßen, aber getrennt voneinander evaluiert werden. Die dynamischen Verteilungsmöglichkeiten werden im Zusammenhang mit Lernmanagementsystemen in der Evaluationsstudie evaluiert.

## 7 Konsequenzen für die Evaluationsstudie

In dem ersten Teil dieses Kapitels wurden Definitionsversuche zu den Begriffen *Multimedia* und *Neue Medien* vorgestellt. In Abgrenzung zu diesen wurde der Versuch einer für die Arbeit gültigen Begriffsdefinition vorgestellt. Diese stellt die Interaktion des Anwenders mit dem System in den Mittelpunkt. Der Anwender hat durch die neuen Medien viele Interaktionsmöglichkeiten und konstruiert sein Wissen durch Aktivität.

Die Diskussion um den Einsatz multimedialer Lernsoftware wurde kritisch dargestellt. Für die Evaluationsstudie dieser Arbeit wurden die folgenden Erkenntnisse und Konsequenzen abgeleitet:

- Der gezielte Einsatz und die Evaluation multimedialer Lernsoftware müssen organisiert werden. Es bedarf der Entwicklung eines Einsatzkonzeptes für *Statistik interaktiv komplett*. Die Untersuchung verschiedener Lerntheorien hat gezeigt, dass der Konstruktivismus am besten geeignet ist, klassifizierten Defiziten in der Statistikausbildung entgegenzuwirken. Daher werden konstruktivistische Prinzipien in die Lernsoftware und das Einsatzkonzept integriert. Die Evaluationsstudie hat die Aufgabe, diese Umsetzungen hinsichtlich ihrer Zielerreichung zu überprüfen. Dabei muss untersucht werden, ob die Komponenten der Lernsoftware oder deren Kombination und Einsatz zur Zielerreichung (oder „Nichterreichung“) führen.
- Des Weiteren wurde herausgestellt, dass an den Hochschulen neue Strukturen für den Einsatz multimedialer Lernsoftware benötigt werden. Im Rahmen der Evaluati-

onsstudie muss überprüft werden, ob z.B. die Einführung einer neuen Prüfungsform zur Erreichung definierter Ziele führt.

- Die Diskussion hat zudem gezeigt, dass die neuen Medien eine neue Art der didaktischen Aufbereitung der Inhalte verlangen. Die Evaluationsstudie muss die damit verbundenen Probleme herausstellen und in *Statistik interaktiv komplett* implementierte Lösungen bewerten.
- Es wurden verschiedene Ansätze zur Evaluation im Allgemeinen und speziell in Bezug auf multimediale Lernsoftware diskutiert. Die Evaluationsstudie muss so konzipiert und durchgeführt werden, dass sie Scrivens Charakterisierung der *Echten Evaluation* genügt. Zudem müssen bei der Konzeption und Durchführung der Evaluationsstudie die *Standards für Evaluationen* von Beywl berücksichtigt werden. Von dem vorgestellten Vierebenenmodell von Kirkpatrick sind die drei Ebenen Reaction, Learning und Behavior für die Evaluationsstudie relevant.

Spezielle Ansätze zur Evaluation multimedialer Lernsoftware wurden diskutiert. Es wurde die Entscheidung abgeleitet, in der Evaluationsstudie phasenweise zu evaluieren. Die Evaluation soll dabei als Bewertung verstanden werden. Es findet die Kombination qualitativer und quantitativer Evaluationsverfahren statt.

Im Sinne der vorgestellten evolutionären Gestaltungsansätze hat die Evaluationsstudie die übergeordnete Aufgaben, die folgenden Aspekte experimentell zu erforschen:<sup>202</sup>

- Wie kann durch eine lernförderliche Infrastruktur der Aufwand für die Lehrenden reduziert werden?
- Wie können Projekte so evaluiert werden dass Erfahrungen und Ergebnisse einen Qualitätsnachweis und vergleichbare Anhaltspunkte für eine Verallgemeinerung liefern?
- Wie können sich durch neue Medien neue Qualitäten ergeben?

Dabei wird die Evaluation mittels Kriterienkatalogen ausgeschlossen. Die Evaluation verfolgt vielmehr das Ziel, individuelle Kriterien für konzipierte Einsätze abzuleiten und hinsichtlich ihrer Eignung empirisch zu untersuchen.

Es wurde deutlich, dass der didaktische, organisierte und in lerntheoretische Konzepte eingebettete Multimediaeinsatz so individuell ist, dass bei einem Vergleich verschiedener Projekte Schwierigkeiten aufgrund fehlender Abgrenzungskriterien auftreten. Es scheint keine Möglichkeit zu geben, einen effizienten Vergleich der auf dem Hochschulmarkt befindlichen Multimediaprojekte durchzuführen. Da es keine allgemeinen Kriterien gibt, die eine klare Abgrenzung aller Multimediaeinsätze in der Lehre und eine Bewertung des aktuellen Standes ermöglichen, wird kein Vergleich zu anderen Evaluationsstudien durchgeführt werden.

---

<sup>202</sup>Diese wurden dem Leser oben als evolutionäre Gestaltungsansätze vorgestellt.

## Teil III

# Prototyping - *Statistik interaktiv komplett*

## 8 Phasenweise Entwicklung multimedialer Lernsoftware

Der Schwerpunkt dieses Teils liegt auf der Entwicklung multimedialer Lernsoftware für die Statistikausbildung. In Teil I wurden die in der Statistikausbildung existierenden Defizite identifiziert und klassifiziert. Als mögliche Alternative, diesen entgegenzuwirken, wurden in Teil II die neuen Medien vorgestellt und kritisch betrachtet. Eine Konsequenz dieser Betrachtungen ist, dass der Entwicklungsprozess multimedialer Lernsoftware und das Produkt selbst bei der Bekämpfung der Defizite in der Statistikausbildung eine entscheidende Rolle spielen. Es wurde die These aufgestellt, dass neue Medien nur dann eine Chance für die Statistikausbildung darstellen, wenn sie sinnvoll entwickelt werden. Welche Aspekte bei dieser Entwicklung und Gestaltung zu beachten sind, wird im Folgenden untersucht. Es wird reflektierend über die phasenweise Entwicklung von *Statistik interaktiv komplett* und eines Einsatzkonzeptes berichtet.

Zunächst wird die Methode des Prototypings in das Gebiet "Softwareengineering" eingeordnet. Die theoretischen Betrachtungen werden auf *Statistik interaktiv komplett* übertragen. Die parallele und phasenweise Entwicklung von Lernmodulen und eines Einsatzkonzeptes wird exemplarisch anhand des Statistiklabors demonstriert.

Der zweite Teil dieses Kapitels legt den Schwerpunkt auf die Beurteilung der entwickelten Lernmodule. Dabei steht der Konstruktivismus im Mittelpunkt. Es wird anhand von Beispielen überprüft, inwiefern konstruktivistische Ideen in der Lernsoftware *Statistik interaktiv komplett* und in dem Einsatzkonzept umgesetzt wurden. Die Ausführungen belegen die These dieser Arbeit, dass konstruktivistische Einsatzkonzepte benötigt werden und das hier entwickelte Einsatzkonzept diesen Anforderungen folgt. Die exemplarische Überprüfung dient zudem dazu, dem Leser einzelne Komponenten der entwickelten Lernsoftware und des Einsatzkonzeptes vorzustellen. Es schließt sich eine kritische Reflexion über den Verlauf des Forschungsprojektes *Neue Statistik* an.

### 8.1 Prototyping als Methode der Softwareentwicklung

Es wird untersucht, ob Elemente der Softwareentwicklung geeignet sind, auf die Entwicklung multimedialer Lernsoftware angewendet zu werden. Es wird herausgearbeitet, dass ein dynamischer Ansatz benötigt wird. Auf Basis des evolutionären Gestaltungsansatzes des Softwareengineering wird ein Vorgehensmodell zum Prototyping multimedialer Lernsoftware entwickelt. Dieses wird auf die Entwicklung der Lernsoftware *Statistik interaktiv komplett* angewendet.

#### 8.1.1 Der evolutionäre Entwicklungsansatz

Beginnen wir mit einem Zitat Sommervilles, einem bedeutenden Vertreter des Softwareengineering, der zu bedenken gibt:

„Es gibt keine einfachen Lösungen für die Probleme der Softwareentwicklung, und wir brauchen ein breites Spektrum an Werkzeugen und Verfahren, um die in dieser Wissenschaft bestehenden Probleme zu lösen.“<sup>203</sup>

Dass dieses Zitat auf die Entwicklung der multimedialen Lernsoftware *Statistik interaktiv komplett* zutrifft, wird dem Leser im Folgenden demonstriert. Die für diese Arbeit relevanten Werkzeuge und Verfahren sind das Architekturmodell, das Vorgehensmodell, das phasenweise Prototyping und die evolutionäre Gestaltung. Diese werden im Folgenden genauer betrachtet.<sup>204</sup> Zuvor einige Worte

---

<sup>203</sup>[Sommerville2001], S. 13.

<sup>204</sup>Der Leser sei darauf hingewiesen, dass diese lediglich einen kleinen Teil des Softwareengineering ausmachen.

zum Softwareengineering an sich.

Mit dem Begriff „Software Engineering“ oder „Softwareentwicklung“ verbinden viele die Herstellung eines Computerprogramms. Diese eng gefasste Definition ist oft nicht zutreffend, würde auch unsere Tätigkeiten nicht treffen, da die Entwicklung einer multimedialen Lernumgebung mit vielen verschiedenen Aspekten zusammenhängt. Größtenteils haben diese mit Konzeption zu tun. Dies impliziert der Begriff „Lernumgebung“. Er eignet sich daher besser als der Begriff „Lernsoftware“.

Sommerville bezeichnet die Sicherstellung der Benutzerbedürfnisse als eine der wichtigsten Aufgaben des Softwareengineering.<sup>205</sup> Dieser Aspekt scheint besonders in unserem Zusammenhang von Bedeutung zu sein. Der Benutzer ist sowohl der Studierende, der mit Hilfe von *Statistik interaktiv komplett* lernt, als auch der Dozent, der die multimediale Lernumgebung in der Lehre einsetzt. Betrachtet man traditionelle Ansätze des Softwareengineering, wie z.B. ein Phasenmodell Boehms<sup>206</sup>, so stellt man fest, dass statische Modelle wenig geeignet sind, die Bedürfnisse der Benutzer identifizieren zu können. Die starren Strukturen statischer Modelle sehen keine Rückkopplungen während des Entwicklungsprozesses vor. Es wird die implizite Annahme getroffen, dass die Benutzerbedürfnisse und somit Anforderungen an die zu entwickelnde Software a priori vollständig feststehen und sich nicht oder nur sehr langfristig verändern.<sup>207</sup> Bei der Entwicklung multimedialer Lernumgebungen stehen die Bedürfnisse der Studierenden und Dozenten<sup>208</sup> nicht von vornherein fest. Sie entstehen während des Entwicklungsprozesses. Es empfiehlt sich daher die Auseinandersetzung mit dem Ansatz der evolutionären Softwareentwicklung. Die schrittweise Entwicklung der multimedialen Lernsoftware, die im Rahmen dieser Arbeit vorgenommen wird, würde von Sommerville als „evolutionäre Entwicklung“ bezeichnet werden.

„Dieser Ansatz überlappt die Spezifikation, die Entwicklung und die Validierung. Mit Hilfe abstrakter Spezifikationen wird schnell ein erstes System entwickelt. Dann wird es mit Rückmeldung des Kunden<sup>209</sup> verbessert, bis ein System entsteht, das die Bedürfnisse des Kunden befriedigt. Danach kann das System ausgeliefert werden. Statt dessen kann es auch mit einem strukturierten Ansatz neu implementiert werden, um ein System zu erhalten, das robuster und besser zu warten ist.“<sup>210</sup>

Um die in diesem Ansatz enthaltenen und ermöglichten Rückkopplungen hervorzuheben und zu begründen, warum sie bei der Entwicklung einer multimedialen Lernumgebung von Bedeutung sind, wird im Folgenden ein evolutionäres Vorgehensmodell entwickelt. Dies geschieht in direkter Abgrenzung zu einem traditionellen, statischen Vorgehensmodell des Prototypings.

Der evolutionäre Entwicklungsansatz unterteilt den Entwicklungsprozess in einzelne Phasen. Dies haben die traditionellen Ansätze zwar auch geleistet, ließen aber keine Rückkopplungen zu. Bei der phasenweisen Entwicklung können in jeder Phase Verbesserungsvorschläge und Erkenntnisse der vorangehenden Phase integriert werden. Durch dieses Vorgehen kann die Qualität<sup>211</sup> der entstehenden Lernsoftware verbessert werden. Identifizierte Benutzerbedürfnisse können direkt berücksichtigt werden.

Die in den einzelnen Phasen entstehenden Lernmodule und Komponenten werden als Prototypen, der ganze Entwicklungsprozess als *Prototyping* bezeichnet. Spitta unterscheidet verschiedene Arten von Prototyping, unter denen sich auch der eben beschriebene Ansatz befindet.<sup>212</sup>

---

<sup>205</sup> Vgl. [Sommerville2001], S. 183.

<sup>206</sup> Vgl. [Spitta1989], S. 26.

<sup>207</sup> Vgl. [Spitta1989], S. 27.

<sup>208</sup> Dazu kommt, dass diese eine hohe Variabilität aufweisen - nicht jeder Studierende/Dozent hat dieselben Bedürfnisse.

<sup>209</sup> Der Kunde ist hier der Student und der Dozent, der die Lernsoftware einsetzt.

<sup>210</sup> [Sommerville2001], S.24.

<sup>211</sup> Qualität kann durch den schnellen und effektiven Verlauf von Lernprozessen entstehen.

<sup>212</sup> Vgl. [Spitta1989], S. 5.

- Exploratives Prototyping, welches den Prototypen als Kommunikationsmedium einsetzt, um die relevanten Anforderungen der Benutzer zu ermitteln.
- Evolutionäres Prototyping, das die eben beschriebene stufenweise Entwicklung vorsieht.

Betrachten wir das evolutionäre Prototyping genauer. Die Unterscheidung zwischen der eigentlichen Entwicklung der Lernsoftware und der Entwicklung der Prototypen ist nicht voneinander zu trennen. Aus den schrittweise zu entwickelnden Prototypen ergibt sich im letzten Schritt die Lernsoftware. Der erste Prototyp könnte daher als Anfangsversion, der letzte als Endversion der Software bezeichnet werden. Die schrittweise Entwicklung verschiedener Prototypen soll zur Validierung der Benutzerbedürfnisse führen. Am Ende des Prototypings entsteht idealerweise ein Produkt, das den Bedürfnissen der Benutzer so gut wie möglich gerecht wird.

Nach dieser allgemeinen Einordnung erfolgt die Entwicklung eines geeigneten Vorgehensmodells für das Prototyping multimedialer Lernsoftware und ein dazugehöriger Architekturf Entwurf für *Statistik interaktiv komplett*.

### 8.1.2 Das gewählte Vorgehensmodell

In diesem Abschnitt wird ein Vorgehensmodell zum Prototyping multimedialer Lernsoftware entwickelt, das auf Basis des im vorigen Abschnitt beschriebenen Ansatzes der evolutionären Softwareentwicklung entsteht. Es wird eine notwendige Modifizierung des Ansatzes von Sommerville vorgeschlagen, die speziell für die Entwicklung von Lernsoftware Gültigkeit hat. Das entstehende Vorgehensmodell stellt die einzelnen Schritte der Entwicklung einer multimedialen Lernumgebung im Überblick dar. Es enthält strukturierte Ansätze zum Vorgehen, gibt Ratschläge und Anleitungen für ähnliche Entwicklungsvorhaben.

Vorgehensmodelle dienen dazu, einen bestimmten Prozess abstrakt und vereinfacht zu beschreiben, so dass Dritte auf Basis des Vorgehensmodells angeleitet vorgehen können und Fehler vermeiden. Jedes Vorgehensmodell zeichnet sich dadurch aus, dass es den zu beschreibenden Prozess aus einer bestimmten Perspektive betrachtet. Das im Folgenden beschriebene Vorgehensmodell für die Entwicklung multimedialer Lernsoftware ist in kritischer Abgrenzung zu einem statischen traditionellen Vorgehensmodell entstanden, welches die folgende Gestalt annimmt.<sup>213</sup>

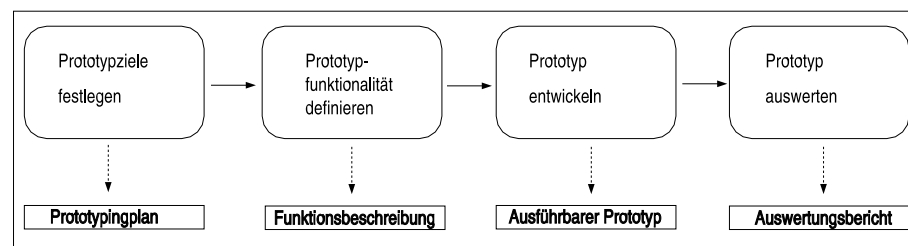


Abbildung 37: Vorgehensmodell zum Prototyping von Sommerville

Die folgenden Ausführungen demonstrieren dem Leser die Notwendigkeit dynamischer Strukturen, wie sie der Ansatz der evolutionären Softwareentwicklung vorschlägt. Die folgende Abbildung zeigt das im Rahmen dieser Arbeit entwickelte Vorgehensmodell zum Prototyping multimedialer Lernsoftware:

<sup>213</sup>Vgl. [Sommerville2001], S. 183.

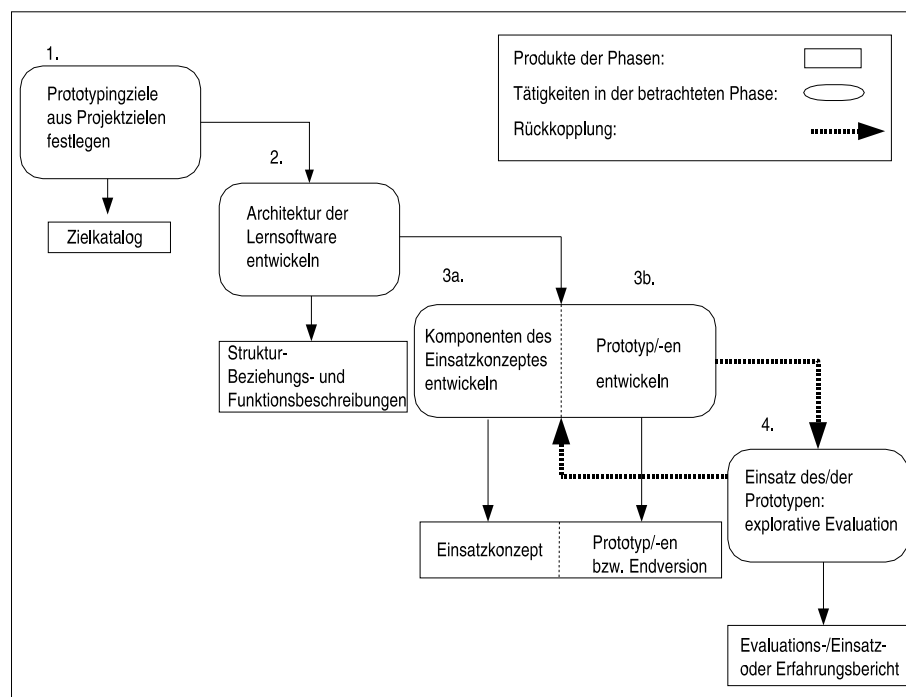


Abbildung 38: Vorgehensmodell zum Prototyping

Zunächst werden die einzelnen Schritte des vorgeschlagenen Vorgehensmodells im Vergleich zu dem statischen Vorgehensmodell betrachtet. Anschließend wird eine geringfügige Modifikation dieses Modells vorgenommen. Diese betrifft keine Phasen, sondern nur Rückkopplungen.

In der **1. Phase** werden die allgemeinen Ziele des Forschungsprojektes festgelegt und herangezogen, um die Ziele des Prototypings festzulegen. Es sollte während der Entwicklung in festen zeitlichen Abständen eine wechselseitige Abstimmung der Entwicklung und Ziele des Prototypings stattfinden. In dem Projekt *Neue Statistik* sind die Prototypingziele und Projektziele<sup>214</sup> ähnlich. Dies liegt nahe, da die zu entwickelnde multimediale Lernsoftware zu einem Großteil für die Umsetzung der Ziele verantwortlich ist. Die übrigen Ziele sollen durch die entwickelten Einsatzkonzepte realisiert werden.

Im Gegensatz zu Sommerville wird auf dieser Ebene nicht zwischen den Zielen einzelner Prototypen unterschieden. Es werden Ziele formuliert, die für den gesamten Entwicklungsprozess gelten, die Prototypingziele. Später bilden sich die untergeordneten Ziele der einzelnen Prototypen heraus.<sup>215</sup> Diese sind als Konsequenzen zu verstehen, die sich aufgrund des Einsatzes/der Evaluation des voranstehenden Prototypen ergeben.

In der **2. Phase** wird im Gegensatz zu Sommerville über die Architektur (Struktur-, Beziehungs- und Funktionsbeschreibungen) der multimedialen Lernsoftware entschieden. Dies betrifft Entscheidungen über die Gestaltung und den Aufbau der Lernsoftware. Es wird festgelegt, dass sich Lernsoftware aus verschiedenen Lernmodulen zusammensetzt: *Statistik interaktiv komplett* wird modular gestaltet. Die Module bestehen wiederum aus Komponenten und bilden zusammen eine in sich geschlossene Einheit. Die Architektur soll ermöglichen, dass Studierende sowohl Teilbereiche, als auch die Statistik als Disziplin verstehen. In diesem Schritt wird auch über die Rollen, Aufgaben, Inhalte der Lernmodule und Komponenten entschieden.

<sup>214</sup> Vgl. 11.

<sup>215</sup> Vgl. 8.3.



Die **3. Phase** widmet Sommerville der Entwicklung des Prototypen. Wir ergänzen die Phase durch die Entwicklung eines Einsatzkonzeptes. Dies folgt der These, dass Einsatz und Entwicklung sich wechselseitig bestimmen und nicht getrennt voneinander betrachtet werden können. Die Eigenschaften eines Einsatzkonzeptes werden im Folgenden detailliert betrachtet, um dem Leser seine Aufgaben, Ideen und Notwendigkeit zu vermitteln:

#### **Definition eines Einsatzkonzeptes**

Ein Einsatzkonzept legt fest, wie die Lernmodule und Komponenten multimedialer Lernsoftware in der Statistikausbildung eingesetzt werden sollen. Dies geschieht in enger Abhängigkeit zu den definierten Zielen des Einsatzes. Ein Einsatzkonzept kann durch die folgenden Größen beschrieben werden:

- Materialien
- Umgang mit den Materialien

Ein Einsatzkonzept soll die Komponenten der Lernmodule so kombinieren, dass den Studierenden bestimmte statistische Kompetenzen vermittelt werden. Aus den identifizierten Defiziten der Statistikausbildung wurden die folgenden erwünschten Kompetenzen der Studierenden abgeleitet:

- Verständnis für statistische Methoden - Methodenverständnis
- Verständnis für gezielten (und richtigen) Einsatz statistischer Methoden zur Problemlösung - Anwendungsverständnis
- Verständnis für Zusammenhänge und Abhängigkeiten zwischen den Methoden
- Verständnis für mathematische Konzepte und schrittweise Beweisführung
- Fähigkeit, statistische Reports zur Dokumentation des statistischen Arbeitens zu verfassen
- Souveräner Umgang mit neuen Medien und Verständnis für die notwendige Zusammengehörigkeit von "Computational" und "Statistics"
- Fähigkeit, organisiert und selektiv zu lernen, Prioritäten beim Lernen zu setzen
- Erster Überblick über Anwendungsfelder, Relevanz und Rolle der Statistik in der Realität

Diese Kompetenzen sollen durch die Kombination der Lernmodule / Komponenten und der konstruktivistischen Einsatzkonzepte erreicht werden. Dies wird in Teil IV detailliert erläutert. Bei dem folgenden Beispiel handelt es sich um einen Auszug aus Teil IV der vorliegenden Arbeit.

#### **Verständnis für statistische Methoden - Methodenverständnis**

Die Theoriekomponenten werden in der Vorlesung zur Vermittlung statistischer Methoden eingesetzt. Sie werden ergänzt durch Vorlesungsfolien/Skript, die die statistischen Methoden zum Teil ausführlicher darstellen. Zum einen ist es notwendig, die Theoriekomponenten zu ergänzen, da diese knapp gefasst sind.<sup>216</sup> Zum anderen können die Ergänzungen als Individualisierung betrachtet werden: Dozenten ergänzen Inhalte, die ihnen persönlich wichtig sind, um bestimmte statistische Inhalte zu vermitteln. Des Weiteren dienen die Theoriekomponenten der Nachbereitung der Vorlesungen durch die Studierenden. Auch Inhalte, wie zum Beispiel die Erhebungsverfahren, die in der Vorlesung nur kurz angesprochen werden, können sich die Studierenden eigenständig durch die Theoriekomponenten aneignen.

Das Glossar wird sowohl in der Vorlesung als auch in der Nachbereitung eingesetzt. Statistische Methoden oder Teile werden knapp, vergleichbar mit Lexikoneinträgen, beschrieben. Vermittelt die Theoriekomponente die Klassierung einer Häufigkeitstabelle, kann im Glossar *Häufigkeitstabelle klassiert* nachgeschlagen werden. Die im Glossar beschriebenen Begriffe sind untereinander

<sup>216</sup>Es handelt sich um ein Gestaltungsproblem. Im Projekt *Neue Statistik* hat sich das *One-page-principle* durchgesetzt. Dieses führte zu einer zu knappen Gestaltung der Theoriekomponenten. In der kritischen Reflektion am Ende dieses Kapitels wird dieser Aspekt ausführlich erläutert.

verlinkt. Es entsteht ein Überblick, welche Begriffe bei bestimmten statistischen Methoden eine Rolle spielen.

Die praktische Anwendung der statistischen Methoden erfolgt im Statistiklabor. So kann z.B. die klassierte Häufigkeitstabelle grafisch durch das Histogramm abgebildet werden. Man kann unterschiedliche Klassenbreiten wählen und deren Auswirkung auf die grafische Darstellung der Dichte der Daten illustrieren. Dazu kann der Grafik-Wizard eingesetzt werden, wenn der Studierende sich allein auf die Ideen konzentrieren soll. Es besteht außerdem die Möglichkeit, den R-Kalkulator einzuführen und durch die Anwendung von Bibliotheken erste Kontakte mit Programmiersprachen herzustellen. Anhand von Minute Papers wird überprüft, ob die vermittelte statistische Methode verstanden wurde.

Die Ausführungen zeigen, dass das Einsatzkonzept die Kombination der multimedialen Komponenten organisiert.

Die **vierte Phase** besteht in dem Einsatz und der Evaluation des Prototypen in der Lehre auf Basis des parallel entwickelten Einsatzkonzeptes. Es entsteht ein Bericht, der über die auftretenden Probleme des entwickelten Prototypen und Einsatzkonzeptes informiert. Die Besonderheit des Vorgehensmodells besteht darin, dass eine Rotationsbewegung einsetzt. Die identifizierten Schwachstellen werden im nächsten Entwicklungsschritt berücksichtigt. Dies betrifft sowohl den Prototypen als auch das Einsatzkonzept. Es findet somit parallel der Versuch einer schrittweisen Optimierung beider statt.

Die Gegenüberstellung sollte folgende Aspekte verdeutlicht haben. Es besteht die Notwendigkeit eines dynamischen Vorgehensmodells, das zyklische Elemente enthält. Traditionelle, starre Vorgehensmodelle der Softwareentwicklung sind für die Entwicklung multimedialer Lernsoftware ungeeignet. Sommerville greift den Ansatz des *Evolutionary Prototyping* in seinen Überlegungen auf und entwickelt das folgende Vorgehensmodell<sup>217</sup>:

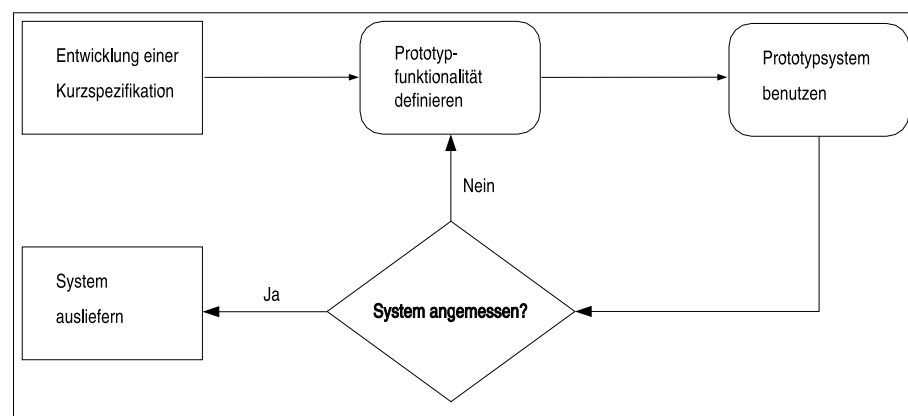


Abbildung 39: Evolutionary Prototyping nach Sommerville

Die Parallelen zu dem oben vorgestellten Vorgehensmodell sind deutlich erkennbar. Der entscheidende Unterschied zu unserer Vorgehensweise besteht darin, dass Sommerville auch in diesem Ansatz keine Einsatzkonzepte in seine Überlegungen integriert. Das in dieser Arbeit vorgestellte Vorgehensmodell ermöglicht die Entwicklung modifizierbarer Lernsoftware, die adjustierbar hinsichtlich der Bedürfnisse der Dozenten und Studierenden ist.

<sup>217</sup> Vgl. [Sommerville2001], S. 186.

„Das Ziel beim Evolutionary Prototyping besteht darin, ein einsatzfähiges System für Endbenutzer zu entwickeln.“<sup>218</sup>

Einsatzfähig soll es sein, über die Gestaltung des Einsatzes sagt das Modell allerdings nichts aus. Dies scheint eine Besonderheit multimedialer Lernsoftware/-systeme zu sein, die deshalb in unserem Modell der besonderen Berücksichtigung bedarf. Übertragen wir den evolutionären Ansatz Sommervilles auf unser Prototyping, so muss der folgende Aspekt herausgestellt werden. Im Softwareengineering wird unterschieden zwischen der Validierung und der Verifikation eines Software-systems.

„Die Verifikation ist die Überprüfung, dass ein Programm mit der entsprechenden Spezifikation übereinstimmt. (...)

Die Validierung soll nicht die Übereinstimmung mit einer Spezifikation überprüfen, sondern demonstrieren, dass das Programm für den vorgesehenen Zweck geeignet ist.“<sup>219</sup>

Die Schwierigkeit besteht darin, dass für den ersten Prototypen keine klar definierten Spezifikationen existieren, da die genauen Anforderungen erst während des Einsatzes herausgearbeitet werden. Dies macht eine Verifikation zu Beginn unmöglich. In den folgenden Entwicklungsphasen kann sie durchgeführt werden. Auch eine Validierung der Lernsoftware kann nur bedingt vorgenommen werden, da sie stark individuell geprägt ist. Die Bedürfnisse der Benutzer, die in den Entwicklungsprozess eingebunden sind, können z.B. erfüllt sein, was nicht bedeutet, dass dies auf alle möglichen Benutzer zutrifft.

#### **Zusammenfassung:**

- Die Entwicklung multimedialer Lernsoftware wird nach einem Vorgehensmodell entwickelt, um den Entwicklungsprozess zu organisieren und die Ziele der Entwicklung nicht aus den Augen zu verlieren.
- Der evolutionäre Entwicklungsansatz, d.h. die schrittweise Entwicklung und Verbesserung von Prototypen, kann dazu beitragen, die Bedürfnisse der Benutzer gezielt zu ermitteln und zu erfüllen. Die Entwicklung ist adjustierbar hinsichtlich der Benutzerbedürfnisse. Außerdem können Projektpartner Umsetzungsideen präsentieren und zur Diskussion stellen, d.h. die Prototypen dienen als Kommunikationsgrundlage.
- Die Anwendung des in dieser Arbeit vorgeschlagenen Vorgehensmodells ermöglicht die Umsetzung des evolutionären Entwicklungsansatzes. Dieser wird um die parallele Entwicklung eines Einsatzkonzeptes erweitert.
- Die parallele Entwicklung von Prototypen und dazugehörigen Einsatzkonzepten hat das Ziel, die Lernsoftware einsatzgerecht zu entwickeln, d.h. einsatzfähige multimediale Lernsoftware entstehen zu lassen.

Nach diesen Ausführungen zum Vorgehensmodell werden die einzelnen Phasen genauer betrachtet. Im folgenden Abschnitt erfolgt die Festlegung der Prototypingziele aus den Zielen des Gesamtprojektes.

---

<sup>218</sup>[Sommerville2001], S. 185.

<sup>219</sup>[Sommerville2001], S. 186.

### 8.1.3 Projektziele und Prototypingziele

Das für diese Arbeit entwickelte Vorgehensmodell<sup>220</sup> zum Prototyping multimedialer Lernsoftware sieht in der ersten Phase die Festlegung der Prototypingziele aus den Projektzielen vor. Die **Projektziele** (übergeordnete Ziele des Forschungsprojektes *Neue Statistik*) sind im Folgenden definiert:

1. Entwicklung einer multimedialen Lernumgebung zum ganzheitlichen Erlernen von Statistik
2. Nutzung der Möglichkeiten der neuen Medien, um die Statistikausbildung zu gestalten. Die Gestaltung soll Defiziten entgegenwirken und statistische Kompetenzen vermitteln. (Die Zieleigenschaften der Studierenden wurden dem Leser bereits vorgestellt.)
3. Entwicklung einer Lernumgebung, die flächendeckend an Universitäten in ganz Deutschland eingesetzt werden kann.

Aus diesen werden die **Ziele für das Prototyping** als Entwicklungsprozess der multimedialen Lernmodule und -komponenten abgeleitet. Diese sollen dazu führen, dass die Projektziele erreicht werden:<sup>221</sup>

- Vermittlung der Statistik als eine Disziplin, die sich zusammensetzt aus statistischem Problemlösen, statistischem Denken und statistischen Methoden  
(⇒ Projektziel 2)
- Gestaltung der Lernsoftware und des Einsatzkonzeptes nach konstruktivistischen Elementen und Beachtung der Defizite in der Statistikausbildung  
(⇒ Projektziel 2)
- Sicherstellung der Modifizierbarkeit der multimedialen Lernumgebung, um dem flächendeckenden und zeitlosen Einsatz gerecht zu werden (⇒ Projektziel 3)
- Berücksichtigung der interdisziplinären Einsatzfähigkeit bei der Gestaltung, da Statistik in vielen Disziplinen gelernt und benötigt wird (⇒ Projektziele 1, 2, 3)

Diese Ziele stellen einen erwünschten Idealzustand des Entwicklungs- und Produktionsprozesses dar. Der Leser wird im Verlauf der Arbeit erkennen, dass die Projektarbeit die Ziele nur bedingt umsetzen konnte.

### 8.1.4 Architekturentwurf

Der Entwurf einer geeigneten Architektur für die multimediale Lernsoftware ist in der zweiten Phase des Vorgehensmodells vorgesehen. Die Architektur spielt eine grundlegende Rolle im Entwicklungsprozess. Sommerville spricht von „Softwarearchitektur“ und bezeichnet diese als „eine stark vereinfachte Darstellung des Systems, die als Diskussionsgrundlage für verschiedene Projektbeteiligte dienen kann.“<sup>222</sup> Die Architektur gibt einen Überblick über die Struktur der Lernsoftware. Sommerville unterscheidet zwischen verschiedenen Architekturmodellen bezüglich des Inhaltes, von denen wir zwei betrachten:<sup>223</sup>

1. Ein *statisches Strukturmodell*, das die Subsysteme oder Komponenten darstellt, die als separate Einheiten entwickelt werden sollen.
2. Diverse *Beziehungsmodelle*, welche die Beziehungen zwischen den Subsystemen, wie zum Beispiel die Datenflüsse, verdeutlichen.

Die Lernmodule stellen bei *Statistik interaktiv komplett* die von Sommerville angesprochenen Subsysteme dar, die zusammen das gesamte System ergeben. Die Architektur der multimedialen Lernsoftware *Statistik interaktiv komplett* wird durch drei Modelle beschrieben:

<sup>220</sup>Vgl. Abb. 38.

<sup>221</sup>Durch die Angaben in den Klammern wird der konkrete Bezug zu den Projektzielen verdeutlicht.

<sup>222</sup>[Sommerville2001], S. 225.

<sup>223</sup>Vgl. [Sommerville2001], S. 227.

- Strukturmodell<sup>224</sup>: Darstellung der Zusammensetzung der Lernsoftware (Themengebiete, Lernmodule, Komponenten)
- Beziehungsmodell für Ebene II<sup>225</sup> (exemplarisch): verschiedene Perspektiven auf die statistische Datenanalyse
- Beziehungsmodell für Ebene III<sup>226</sup> (exemplarisch): zentrale Rolle des Statistiklabors

**Strukturmodell:** Die Lernsoftware ist unterteilt in Themengebiete. Zu jedem werden multimediale Lernmodule entwickelt. Diese bilden gemeinsam die Einheit *Statistik*. Alle Lernmodule sind untereinander durch Verlinkungen verbunden. Jedes Lernmodul setzt sich aus multimedialen Komponenten zusammen. Es sei angemerkt, dass nicht jedes Lernmodul alle aufgeführten Komponenten enthält. Es werden jeweils die Komponenten entwickelt, die zum Transport der jeweiligen statistischen Inhalte benötigt werden.

Die folgende Grafik dient der Veranschaulichung dieses Aufbaus. Es entsteht ein Strukturmodell auf drei Ebenen.<sup>227</sup>

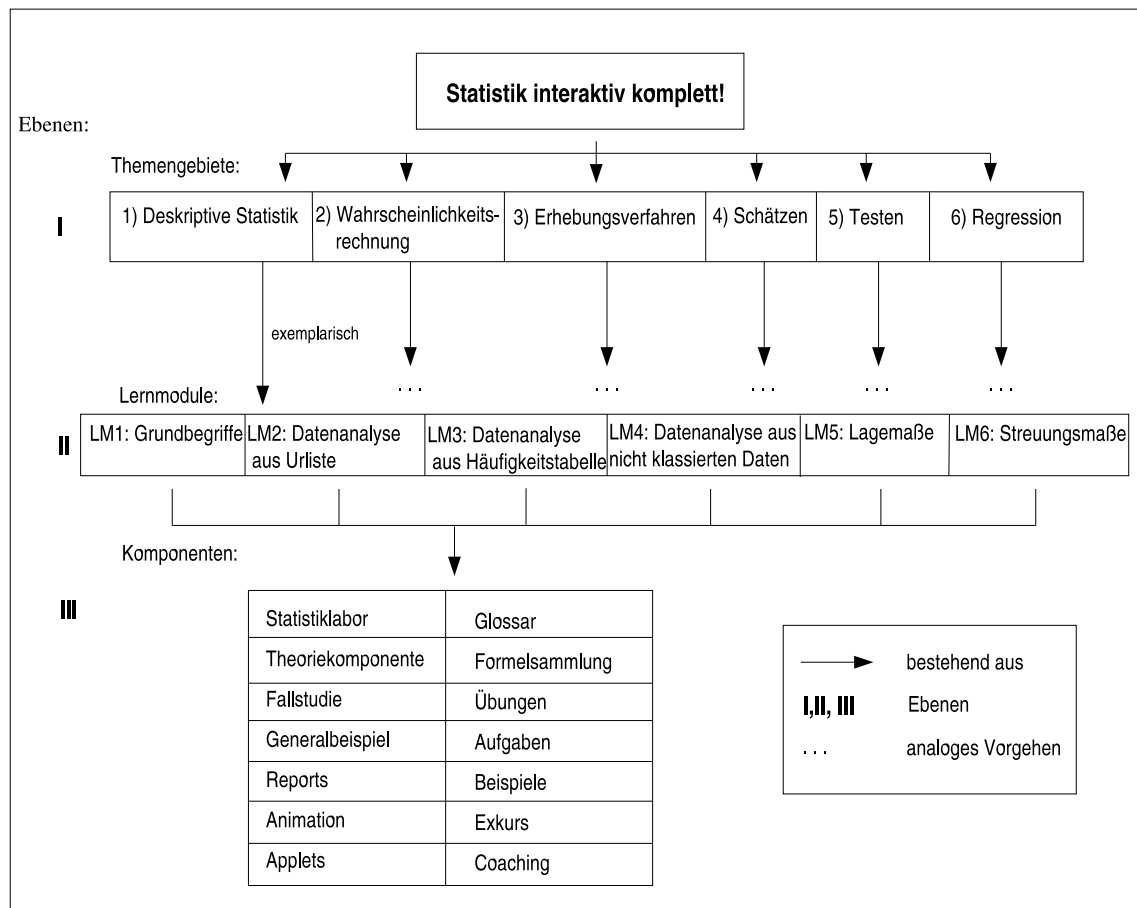


Abbildung 40: Strukturmodell (Ebene I) für die entwickelte Lernsoftware

<sup>224</sup> Vgl. Abb. 40.

<sup>225</sup> Vgl. Abb. 41.

<sup>226</sup> Vgl. Abb. 42.

<sup>227</sup> Inwiefern einzelne Komponenten den oben genannten Zielen folgen, wird in der anschließenden Beschreibung der Komponenten deutlich.

Die beiden folgenden Modelle bilden bestehende Beziehungen zwischen Lernmodulen und Komponenten auf den Ebenen II und III des Strukturmodells exemplarisch ab. Beide zeigen, dass bei der Gestaltung der Architektur und Lernsoftware ein Einsatzkonzept das entscheidende Kriterium war. Sowohl die Struktur als auch die Beziehungen/Datenflüsse zwischen Lernmodulen und Komponenten sind durch die Art und die Ziele des Einsatzes bestimmt. Es werden Thesen aufgestellt, wie bestimmte Ziele erreicht werden können. Diese Thesen bestimmen den Einsatz und die Gestaltung der Module und Komponenten.

Die exemplarische Vorgehensweise verdeutlicht die Überlegungen, die mit dem Entwurf einer Architektur für *Statistik interaktiv komplett* verbunden waren. Die Beziehungen, die zwischen den statistischen Themengebieten auf Ebene I bestehen, haben sich durch den Einsatz einer Lernsoftware nicht verändert. Dies muss kritisch gesehen werden. In dem Forschungsprojekt *Neue Statistik* fand keine Diskussion alternativer Ansätze und Darstellungs-/Vermittlungsformen statt. Ansätze wie z.B. in *Statistics from Scratch*<sup>228</sup> wurden bei der Gestaltung nicht berücksichtigt.

**Beziehungsmodell für Ebene II (exemplarisch):** Die Lernmodule der einzelnen Themengebiete sind nach verschiedenen Kriterien gestaltet und miteinander verknüpft. Die Kriterien sind abhängig von den jeweiligen statistischen Inhalten der Lernmodule. Exemplarisch wird die Zerlegung und Verknüpfung für die folgenden Lernmodule des Themengebietes deskriptive Statistik dargestellt. Die statistische Datenanalyse wird nach dem konstruktivistischen Prinzip *Multiple Perspectives* vermittelt. Der Studierende lernt in vier Lernmodulen, inwiefern deskriptive Datenanalysen durchgeführt werden können. Dabei liegen die zu analysierenden Daten in unterschiedlicher Form vor.

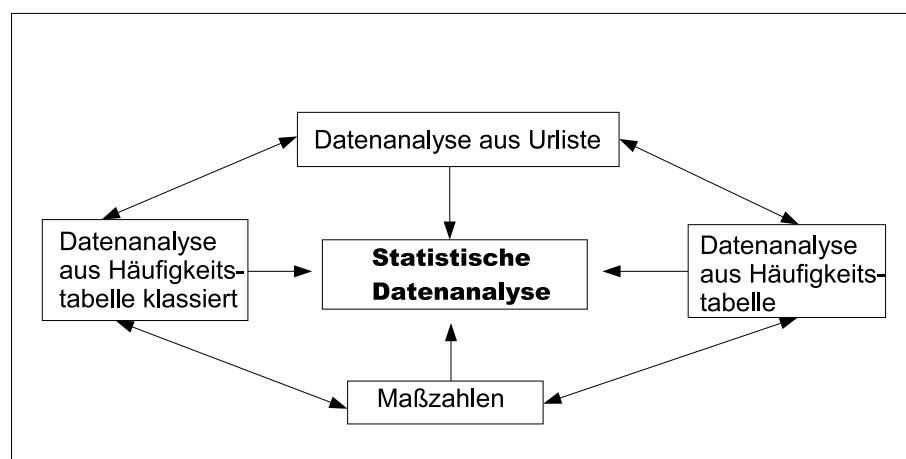


Abbildung 41: Beziehungsmodell Ebene II: Verschiedene Perspektiven auf statistische Datenanalyse

Diese Gestaltung hat das Ziel, statistisches Methodenwissen durch wiederholte Vermittlung zu vertiefen. Die Wissenskonstruktion und -vernetzung wird positiv beeinflusst.

”Multiple perspectives, authentic activities, real-world environments these are just some of the themes that are frequently associated with constructivist learning and teaching.”<sup>229</sup>

Das folgende Beziehungsmodell greift die Aspekte *authentic activities* und *real-world environments* des Zitats auf und zeigt, inwiefern diese die Gestaltung der Komponenten beeinflusst haben.

<sup>228</sup>Vgl. [Bowers1996].

<sup>229</sup>[Murphy1997].

**Beziehungsmodell für Ebene III (exemplarisch):** Zwischen zahlreichen im Strukturmodell dargestellten Komponenten bestehen Beziehungen, fließen Daten. Es wäre unübersichtlich, alle existierenden Datenflüsse in einem Modell abzubilden. Im Zentrum der Gestaltung der multimedialen Komponenten stand das Statistiklabor, da es viele Ziele des Projektes unterstützen kann. Das Statistiklabor wurde während des Entwicklungsprozess so gestaltet, dass es den Aspekten *authentic activities* und *real-world environments* gerecht wird. Die Analyse realer Datensätze, die Erstellung statistischer Reports sowie die Lösung von in Fallstudien aufgeworfenen Problemen finden im Statistiklabor statt. Der Studierende hat im Statistiklabor die Möglichkeit der Aktivität/Interaktivität. Er kann Datenanalysen des Dozenten nachvollziehen und/oder eigene Analysen in Zusammenarbeit mit anderen Studierenden durchführen.

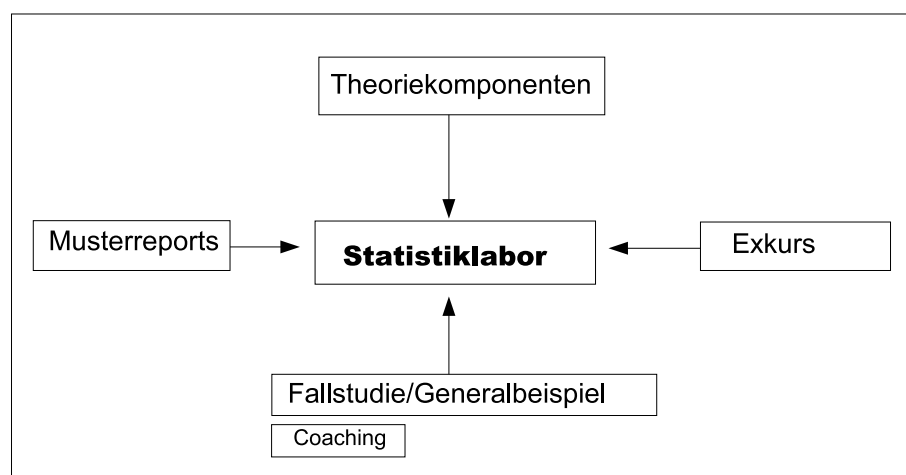


Abbildung 42: Beziehungsmodell Ebene III: Zentrale Rolle des Statistiklabors

Die Verknüpfung der Theoriekomponenten und des Statistiklabors basiert auf der folgenden Überlegung. Die theoretischen Konzepte statistischer Methoden sollten zeitnah durch reale Anwendungsbeispiele verdeutlicht werden, um das Verständnis der Studierenden zu fördern. Die Musterreports demonstrieren, wie statistische Fragestellungen in Reportform durch den Einsatz des Statistiklabors gelöst werden können.

Anhand der Coachingkomponente wird ein Problem des Forschungsprojektes *Neue Statistik* verdeutlicht. Es waren gute Ideen vorhanden, die jedoch schlecht umgesetzt wurden.

Die Coachingkomponente übernimmt das *Guiding* der Studierenden, erklärt z.B. die Schritte des Erstellungsprozesses eines statistischen Reports. Das Beziehungsmodell zeigt, dass die Coachingkomponente nicht auf einer Ebene mit den anderen Komponenten angesiedelt ist. Dies ist nicht auf Gestaltungsaspekte, sondern auf Projektpolitik zurückzuführen. Durch mangelnde Diskussion wurde die Rolle und Notwendigkeit der Coaching-Komponente nicht erkannt bzw. unterschätzt. So wurden nur im Themengebiet *Deskriptive Statistik* Coaching-Komponenten entwickelt und implementiert. Aufgrund der fehlenden Bereitschaft vieler Projektpartner, ihre Lernmodule durch eine Coaching-Komponenten zu erweitern, wurde diese alternativ verarbeitet und eingebunden. Sie erscheint nicht als gleichberechtigte Komponente, sondern als Unterkomponente der Komponente *Fallstudie*. Dies wird in der obigen Abbildung deutlich. Das Problem besteht darin, dass der Studierende keine Möglichkeit hat, von einer anderen Komponente als der Fallstudie auf die Coachingkomponente zuzugreifen. Bearbeitet er z.B. einen Musterreport und erstellt einen eigenen Report im Statistiklabor, hat er keinen direkten Zugriff auf die Coachingkomponente *Statistisches Reportverfassen*. Diese erläutert die Schritte und Ideen der Reporterstellung. Die Unterordnung der Coachingkomponente ist auf das Streben nach einer *Corporate Identity* zurückzuführen. Um eine einheitliche Gestaltung aller Lernmodule zu erreichen, hätte die Coachingkomponente auch in

die Lernmodule anderer Projektpartner eingebunden werden müssen. Da viele diese Notwendigkeit nicht sahen, wurde sie untergeordnet. Dies führte zu einer Schwäche der Lernmodule. Nach dieser kritischen Reflexion wird eine Modifikation des entwickelten Vorgehensmodells hinsichtlich der Architektur diskutiert.

### Notwendige Modifikation des Vorgehensmodells

In dem entwickelten Vorgehensmodell findet durch die evolutionäre Entwicklung eine permanente Verbesserung/Modifikation der Prototypen statt. Die Architektur ist in diesen Prozess nicht eingeschlossen. Es kann passieren, dass zu Projektbeginn eine Architektur entwickelt wird, die zu diesem Zeitpunkt sinnvoll erscheint, aber im Entwicklungsprozess nicht modifiziert wird. Immer wenn im Entwicklungsprozess ein Problem auftaucht, wird dies dem Vorgehensmodell gemäß<sup>230</sup> entweder auf den Prototypen selbst oder das zugrunde liegende Einsatzkonzept zurückgeführt. Ist ein Problem auf die Architektur zurückzuführen, wird dies nicht erkannt. Diese Überlegung hat zu einer Modifikation des entwickelten Vorgehensmodells geführt. Die Architektur wird in den iterativen Entwicklungsprozess der Prototypen einbezogen. Die zweite Phase (*Architekturentwurf*) wird allerdings nicht in die Rückkopplung einbezogen. Dies würde dazu führen, bei jedem neuen Prototypen und Einsatzkonzept die Grundlage anzuzweifeln. Die Architektur sollte nicht als geschlossenes Curriculum kritisch überdacht werden. Die Funktionalitäten der in ihr enthaltenen Module und Komponenten und insbesondere ihre Beziehungen (Datenflüsse) untereinander sollten hinterfragt werden.

Dieses Problem ist in dem Forschungsprojekt *Neue Statistik* aufgetaucht: eine Architektur auf die sich alle Projektpartner einigen konnten, wurde auf einem der ersten Projekttreffen festgelegt. Diese entstand nicht durch einen Kommunikations- und Argumentationsprozess, sondern wurde durch einige wenige Projektpartner ausgearbeitet und von den anderen akzeptiert. Nach knapp zwei Jahren, in denen die einzelnen Projektpartner Lernmodule entwickelten, wurde festgestellt, dass viele lediglich die äußere Struktur der Architektur einhielten, die Funktionalitäten der Komponenten allerdings traditionell auffassten. Lehrbücher wurden auf die einzelnen Komponenten aufgeteilt.<sup>231</sup> Die Möglichkeiten der neuen Medien wurden nicht gezielt genutzt, z.B. wurden keine Laborseiten dort bereitgestellt, wo Interaktivität zur Verdeutlichung statistischer Konzepte beitragen könnte. Wie man hätte vorgehen können, zeigen die Lösungsvorschläge und Umsetzungen des Teils I der vorliegenden Arbeit.<sup>232</sup>

Die Überlegungen haben zu einer Modifikation des Vorgehensmodells geführt:

---

<sup>230</sup>Vgl. Abb. 38

<sup>231</sup>Wie oben kritisiert, wurden nur Lehrbücher mit üblichen Vorgehensweisen zur Vermittlung der Statistik herangezogen. Alternative Ansätze, die z.B. eine problemorientierte Vermittlung statistischer Methoden vorsahen, wurden nicht berücksichtigt.

<sup>232</sup>Vgl. 4.4.



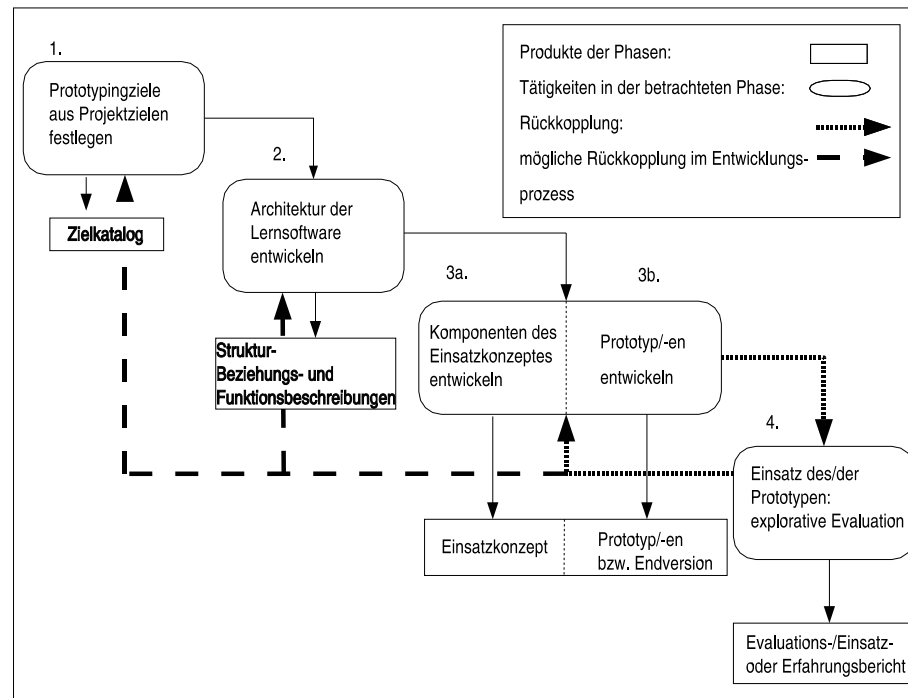


Abbildung 43: Erweitertes Vorgehensmodell

Diese Modifikation des Vorgehensmodells ermöglicht Rückkopplungen bis hin zur ersten Phase, je nach Stärke/Grad des Auslösers. Probleme in der Einsatzphase können also so grundlegend sein, dass die Rückkopplung anders als im zuvor vorgeschlagenen Vorgehensmodell über die Phase des Prototypings hinaus wirkt, auch die Architekturphase umfasst. Welcher Auslöser welche Rückkopplung nach sich zieht, wird exemplarisch bei der Beschreibung der Schwerpunkte und Ziele der Phasen des Prototypings diskutiert.<sup>233</sup>

In dem folgenden Abschnitt werden die Funktionalitäten der multimedialen Komponenten dargestellt. Der Leser bekommt einen Einblick in die phasenweise entwickelte Lernsoftware *Statistik interaktiv komplett*. Es handelt sich um die **Phase 3b: Prototypentwicklung** des Vorgehensmodells. Die Parallelität der Entwicklung des Einsatzkonzeptes (3a) und des Prototypen (3b) wird exemplarisch demonstriert.

## 8.2 Lernmodule von *Statistik interaktiv komplett*

Jedes Lernmodul setzt sich aus Komponenten zusammen. Die Funktionalität, Rolle und Kombinierbarkeit der Komponenten werden beschrieben. Außerdem wird der Zusammenhang zwischen einzelnen Komponenten, konstruktivistischen Prinzipien und Defiziten der Statistikausbildung hergestellt. Es wird demonstriert, wie sich einzelne Komponenten im Entwicklungsprozesses verändert haben. Dies geschieht im Zusammenhang mit der Darstellung der Schwerpunkte und Ziele der drei Phasen des Prototypings.

<sup>233</sup> Vgl. 8.3.

### 8.2.1 Multimediale Lernmodule und Komponenten

**Multimediale Fallstudien** Betrachten wir exemplarisch die Lernmodule zur deskriptiven Statistik: Es wird unterschieden zwischen einer Videofallstudie und einer Fallstudie, die keine Videosequenzen sondern, das Statistiklabor in den Mittelpunkt stellt. Beide haben das Ziel, die einzelnen Lernmodule der deskriptiven Statistik thematisch miteinander zu verbinden. Sie unterscheiden sich inhaltlich wie folgt: Die Basis für die Fallstudie bilden Daten, die an Studierenden erhoben wurden. Die technische Gestaltung der Komponente ermöglicht es dem Dozenten, diese Daten neu zu erheben und in die Fallstudie einzubauen.

Die Videogeschichte besitzt Demonstrationscharakter. Sie basiert auf einem realen Problem aus dem täglichen Leben (Lärmbelästigung durch hohes Verkehrsaufkommen in einem Wohnviertel), das schrittweise durch die Anwendung statistischer Methoden gelöst wird. Die Szenen spiegeln das schrittweise Vorgehen eines Statistikers bei der Problemlösung wider. Dem Studierenden werden die verschiedenen statistischen Methoden durch ihre direkte Anwendung vermittelt. Er soll motiviert werden, diese zu erlernen und zueinander in Verbindung zu setzen. Bei der Bearbeitung bleibt er zunächst passiv.

In der Fallstudie hingegen wird die Aktivität des Studierenden geweckt. Auf vorbereiteten Laborseiten hat er die Möglichkeit, die Daten zu analysieren. Die Gestaltung der Fallstudie folgt dem konstruktivistischen Prinzip der Wissenskonstruktion durch Aktivität bzw. Interaktivität.

**Das Statistiklabor** Das Statistiklabor nimmt eine zentrale Rolle in den Lernmodulen und somit im Prototyping ein. Seine ursprüngliche Rolle, eine Umgebung bereitzustellen, die dem Studierenden Interaktivität in Form der Möglichkeit zur Datenanalyse gibt, wird ausgeweitet. Die vielseitigen Möglichkeiten des Statistiklabors wurden dem Leser in den in Teil I ausgearbeiteten Lösungsvorschlägen demonstriert. Betrachten wir weitere Einsatzmöglichkeiten des Statistiklabors:

- Vorbereitung und Bereitstellung von Laborseiten als Basis für Vorlesungen
- Umgebung zur Bearbeitung von Aufgaben
- Umgebung, die die Demonstration verschiedener Lösungsvarianten ermöglicht
- Umgebung zur Durchführung statistischer Datenanalysen,
- Erstellung statistischer Reports durch eine integrierte Berichterstellungsfunktion
- Umgebung zur Kommunikation (z.B. Materialaustausch) zwischen Studierenden und Dozenten

Der Entwicklungsprozess des Statistiklabors wird in 8.4 beschrieben. In den unterschiedlichen Phasen des Prototypings fanden Verbesserungen der Funktionalitäten des Statistiklabors statt. Durch seinen vielseitigen Charakter besetzt es in seiner Endversion die zentrale Rolle in der Lernsoftware *Statistik interaktiv komplett*.<sup>234</sup>

**Editor und Berichterstellungsfunktion** Es handelt sich um eine Komponente, die im Entwicklungsprozess einer starken Veränderung unterlag. Die Komponente Editor wurde schrittweise durch die sogenannte Berichterstellungsfunktion abgelöst. Da dem Leser dieser Schritt nicht vorenthalten werden soll, erfolgt eine kurze Beschreibung der Komponente Editor. Es handelt sich um einen Texteditor, mit dem eingegebene Texte verändert, gespeichert und gedruckt werden können. Laborobjekte können in diesen durch *copy/paste* übernommen und dokumentiert werden. Ein statistischer Report kann durch die Kombination von Labor (Datenanalyse) und Editor (Dokumentation) erstellt werden: Datenanalyse im Labor, Dokumentation im Editor.

Die Berichterstellungsfunktion, die an die Stelle des Editors tritt, erleichtert die Erstellung statistischer Reports. Sie wird direkt im Statistiklabor aufgerufen, ist also keine eigenständige Komponente mehr. Alle auf der entsprechenden Laborseite vorhandenen Objekte werden bereitgestellt

---

<sup>234</sup>Die Auszeichnung mit dem Hochschulpreis Mediaprix 2003 belegt, dass das Statistiklabor eine wettbewerbsfähige Komponente darstellt.

und können per Mouseclick ausgewählt werden. Durch die Betätigung eines Buttons wird ein Bericht erstellt, in dem der Studierende sein Arbeiten dokumentieren kann. Das Statistiklabor enthält durch die Implementierung der Berichterstellungsfunktion folglich eine Art Reportgenerator, ein Hilfsmittel zur Erstellung statistischer Reports.

**Der statistische Report** Um sowohl den Projektpartnern<sup>235</sup> als auch den Studierenden<sup>236</sup> zu verdeutlichen, was unter einem statistischen Report verstanden wird, wurde die Komponente *Statistischer Report* entwickelt und implementiert. Die in ihr enthaltenen Reports sind inhaltlich eng angelehnt an die für das entsprechende Lernmodul relevante Fallstudie. Es werden Fragen zu ihr aufgeworfen und in Form eines statistischen Reports beantwortet. Dazu werden das Labor und insbesondere die Berichterstellungsfunktion verwendet, um die Reporterstellung zu demonstrieren. Verschiedene Verknüpfungen werden hergestellt:

- zwischen den Komponenten Labor und Fallstudien
- zwischen den Komponenten Report und Labor
- zwischen statistischem Problemlösen und statistischem Reportverfassen

Die Komponente *Statistischer Report* wurde erst nach dem Einsatz des ersten Prototypen entwickelt. Der Einsatz zeigte, dass es nicht ausreicht, theoretisch über statistisches Reportverfassen zu berichten. Die Studierenden benötigen zusätzlich ein Beispiel, um selbstständig ihre ersten Reports erstellen zu können. Dem bereitgestellten Report liegt das konstruktivistische Prinzip des Wissenserwerbs durch Interaktivität zugrunde. Die Laborseiten des Reports können vom Benutzer modifiziert werden.

**Die Theoriekomponenten** Diese Komponenten enthalten die theoretischen Grundlagen (Formeln, Zusammenhänge, Beweise, Herleitungen) der statistischen Methoden der einzelnen Lernmodule. Die Erfahrungen in der Lehre haben gezeigt, dass sich die Studierenden mit der Theoriekomponente viel auseinandersetzen. Die Struktur (und damit auch die Rolle) der Theoriekomponente unterlag im Entwicklungsprozess einer kontroversen Diskussion. Die Projektgruppe Bielefeld vertrat die Position,<sup>237</sup>

- die Texte fein zu untergliedern und
- kleine Lernmodule<sup>238</sup> zu bilden.

Diese Gestaltung folgt zwei Thesen:

**1. These:** Mehrere kleine Lernmodule lassen sich differenzierter miteinander verknüpfen. Zahlreiche Verknüpfungen zeigen dem Studierenden Zusammenhänge zwischen statistischen Methoden auf, führen zu einem höheren Lernerfolg als die einzelne Betrachtung der Methoden mit wenigen Verknüpfungen. Dies folgt dem konstruktivistischen Prinzip der Wissensvernetzung. Die Darstellung demonstriert die Vernetzung der statistischen Gebiete, zeigt Zusammenhänge auf. Die Vernetzung kann vom Studierenden für das eigene aktive Lernen übernommen werden. Ihn auf diese Vernetzung aufmerksam zu machen, ist dem Konstruktivismus gemäß Rolle des Dozenten.

**2. These:** Die wiederkehrenden Strukturen der theoretischen Texte ermöglichen dem Studierenden eine bessere Orientierung. Der Konstruktivismus besagt, dass der Lernende durch bekannte Strukturen organisierter lernen, einen höheren Lernerfolg erzielen kann. Die Abbildung 50 veranschaulicht die Gestaltung der Theoriekomponenten dieser These gemäß.

<sup>235</sup>Dies geschah während des Entwicklungsprozesses.

<sup>236</sup>Dies geschah beim Einsatz der Endversion der multimedialen Lernsoftware *Statistik interaktiv komplett*.

<sup>237</sup>Diese konnte sich nicht durchsetzen, was in der kritischen Reflektion zum Projektablauf *Neue Statistik* erläutert wird.

<sup>238</sup>Lernmodule, die z.B. nur *eine* statistische Methode erklären.

**Die Formelsammlung** Der Einsatz der Prototypen hatte die Entwicklung einer Formelsammlung zur Konsequenz. Diese wurde in der Endversion der Lernsoftware implementiert. Während des Einsatz der Prototypen, die noch keine Formelsammlung enthielten, sah das Einsatzkonzept Folgendes vor: Die Studierenden konnten sich durch die Theorievorlesung und die Hilfestellungen des Dozenten eine Formelsammlung selbst erstellen. Dies folgte dem Prinzip der Wissenskonstruktion durch Aktivität. Die Entwicklung und Implementierung einer Formelsammlung für alle Lernmodule von *Statistik interaktiv komplett* war mit folgenden Fragen verbunden:

- Soll die Formelsammlung eine eigene Komponente darstellen?
- Von welchen Komponenten aus soll die Formelsammlung abrufbar sein?
- Soll die Formelsammlung modifizierbar gestaltet werden?
- Soll die Formelsammlung Umsetzungen der jeweiligen Methoden im Statistiklabor (in R) enthalten?

Es entstand eine Formelsammlung, die alle diese Fragen positiv beantwortet. Sie stellt eine eigene Komponente dar und ist daher von allen anderen Komponenten der Lernsoftware aus abrufbar. Sie ist modifizierbar, kann durch eigene Einträge vom Dozenten ergänzt werden. Die Formelsammlung ist mit Theoriekomponenten, dem Glossar sowie dem Statistiklabor in Form von Hyperlinks verknüpft.

**Das Glossar** Das Glossar stellt ein statistisches Lexikon dar, das dem Studierenden als Nachschlagewerk dienen kann. Es enthält kurze prägnante Erklärungen zu sämtlichen in der Lernsoftware verwendeten statistischen Begriffen. Die Erklärungen sind durch Hyperlinks verknüpft. Begriffe, die zur Erklärung einer statistischen Methode verwendet werden, werden selbst wieder erklärt.

**Applets und Animationen** Die neuen Medien stellen neue Möglichkeiten der Vermittlung statistischer Konzepte durch Visualisierung bereit. Animationen richten sich an den passiven Studierenden und vermitteln einzelne statistische Methoden mit Hilfe von Bild und Ton. Ein Großteil der Applets sind konstruktivistisch gestaltet. Sie erfordern Aktivität seitens des Studierenden. Sie vermitteln statistische Methoden ohne Ton, geben dem Studierenden die Möglichkeit, durch Experimentieren zu lernen.<sup>239</sup>

**Beispiele, Aufgaben und Übungen** Die Namen dieser Komponenten sind selbsterklärend. Zu den Aufgaben stehen innerhalb der Lernsoftware Musterlösungen bereit, die unter Einsatz des Statistiklabors erstellt wurden. Die Übungen sind mit Laborseiten verlinkt, die einen Startpunkt für die Datenanalyse der Studierenden darstellen. Sie enthalten die benötigten Datensätze, um dem Studierenden technische Schwierigkeiten abzunehmen. Wird *Statistik interaktiv komplett* im Rahmen eines Lernmanagementsystems<sup>240</sup> eingesetzt, können die Musterlösungen zeitversetzt bereitgestellt werden. Das Ziel kann darin bestehen, den Studierenden selbst experimentieren zu lassen und ihm zunächst die Möglichkeit zu nehmen, seine Lösungsschritte mit denen der Musterlösung zu vergleichen.

Nach der Beschreibung der einzelnen Komponenten der Lernsoftware wird in dem folgenden Exkurs die Komponente *Fallstudie* detaillierter betrachtet.

### 8.2.2 Exkurs: Konzeption und Aufbau multimedialer Fallstudien

Der Leser bekommt Einblick in die Entwicklung multimedialer Fallstudien. Der Unterschied zwischen verschiedenen Typen multimedialer Fallstudien wird herausgearbeitet. Abstrakte Aspekte der Entwicklung werden anhand konkreter Umsetzungsbeispiele verdeutlicht. Abschließend wird ein Vorgehensmodell auf Basis der Erfahrungen entwickelt. Dieses umfasst die schrittweise Konzeption und Entwicklung multimedialer Fallstudien für Lernumgebungen.

---

<sup>239</sup> Vgl. 10

<sup>240</sup> Vgl. 13.1.2.

Der Komponente *Fallstudie* kommt die Aufgabe zu, einen möglichen Pfad durch die verschiedenen Lernmodule zu legen. Sie stellt Zusammenhänge zwischen Lernmodulen und deren statistischen Methoden her. Indem sie statistische Probleme anspricht, die in den verschiedenen Lernmodulen detailliert behandelt werden, soll sie die Studierenden dazu motivieren, sich mit den verschiedenen Lernmodulen auseinanderzusetzen. Dem Studierenden wird durch die Fallstudie demonstriert, wie einzelne Methoden, die in den Lernmodulen separat vermittelt werden, in Verbindung angewendet werden können.

Im Folgenden wird exemplarisch vorgestellt, wie die Fallstudie zur deskriptiven Statistik<sup>241</sup> konzipiert und erstellt wurde. Sie umfasst die Lernmodule Grundbegriffe, Häufigkeitstabelle, Stabdiagramm, empirische Verteilungsfunktion und Quantile.

Die Fallstudie basiert auf einzelnen Problemstellungen. Durch einen Fragebogen wurden Daten von Studierenden erhoben. Anhand dieser Daten werden Probleme aufgeworfen. Diese werden im Rahmen der Fallstudie, aber auch in weiteren Komponenten der multimedialen Lernsoftware beantwortet.

Grundsätzlich gibt es zwei verschiedene Vorgehensweisen beim Aufbau einer Fallstudie:

- Fallorientiert und nicht linear: In einer fallorientierten, nicht linearen Fallstudie werden verschiedene, voneinander unabhängige Fälle in einer gemeinsamen Umgebung dargestellt. Die einzelnen Fälle können in beliebiger Reihenfolge und unabhängig voneinander gelöst werden. Es könnte zum Beispiel ein Unternehmen betrachtet werden, das unterschiedliche, sich nicht gegenseitig beeinflussende wirtschaftliche Probleme zu lösen hat.
- Szenenorientiert und linear: In einer szenenorientierten, linearen Fallstudie können die einzelnen Szenen nicht unabhängig voneinander betrachtet werden. Eine die Fallstudie verbindende Story wird in einzelnen, aufeinander aufbauenden Szenen dargestellt. In der Regel sind die einzelnen Szenen nur im Kontext der Story verständlich. In jeder Szene wird ein Lösungsschritt vorgenommen, der aus dem der vorangehenden Szene folgt. Es werden auch hier in jeder Szene verschiedene statistische Probleme und Methoden angesprochen. Diese stehen in Verbindung zu den anderen Szenen. Der Nachteil einer szenenorientierten Fallstudie liegt darin, dass von dem Lernenden erwartet wird, dass er einer vorgegebenen Reihenfolge folgt und nicht willkürlich beliebige Szenen betrachten kann.

In der für die deskriptive Statistik entwickelten Fallstudie wird fallorientiert vorgegangen. Bei einer fallorientierten Fallstudie kann sich der Lernende flexibler und selbstständiger verhalten. Er kann die einzelnen Fälle in beliebiger Reihenfolge lösen. Auch die Konzeption einer fallorientierten Fallstudie ist leichter, da nicht jedes zu behandelnde Problem in einen übergreifenden Zusammenhang gebracht werden muss. Jeder Fall hat die folgende Startseite:

---

<sup>241</sup>Zur deskriptiven Statistik existieren zwei Fallstudien. Hier wird diejenige diskutiert, die keine Videosequenzen enthält, sondern das Statistiklabor für Umsetzungen und Visualisierungen einsetzt.

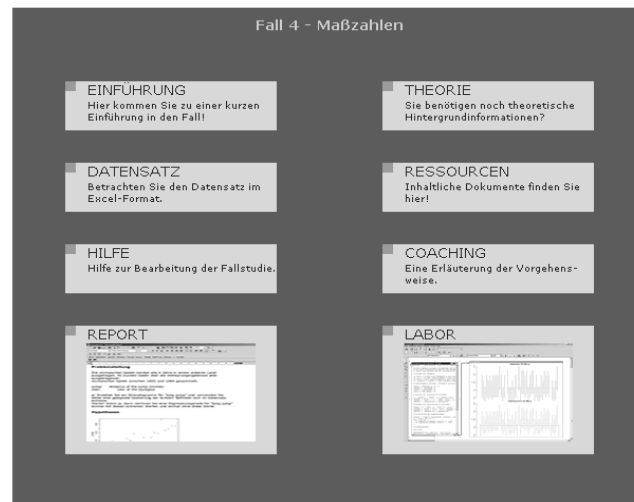


Abbildung 44: Exemplarische Startseite eines Falls

Die Startseiten aller Fälle wurden einheitlich gestaltet, um den Studierenden Orientierung und Wiedererkennung zu ermöglichen. Die acht Felder der Startseite führen den Studierenden in unterschiedliche Gebiete der Fallstudie. Da die einzelnen Fälle in beliebiger Reihenfolge betrachtet werden können, werden keine zusätzlichen Gestaltungselemente benötigt. Jeder Fall enthält die folgenden acht Felder/Bereiche:

- Einführung: Der Studierende wird in die Problemstellung des Falls eingeführt.
- Theorie: Hier werden die statistischen Methoden aufgelistet, die zur Lösung des Falls angewendet werden. Es bestehen Verlinkungen zu den Theoriekomponenten der entsprechenden Lernmodule.
- Datensatz: Es wird der Datensatz, der der Fallstudie zugrunde liegt, bereitgestellt. Dieser kann vom Dozenten ausgetauscht werden.
- Ressourcen: Hier werden zusätzliche Materialien zur Beschreibung des Falls (z.B. Zeitungsartikel) hinterlegt.
- Hilfe: Dem Studierenden werden Hilfestellungen zur Bearbeitung des Falls gegeben. Er wird z.B. auf Laborseiten verwiesen, von denen aus er anfangen kann, Daten zu analysieren.
- Coaching: Hier werden Schritte zur Erstellung des statistischen Reports beschrieben. Zudem wird der Studierende auf Coachingkomponenten verwiesen, die im Zusammenhang mit der Lösung des Falls stehen.<sup>242</sup>
- Labor: Der entsprechende Fall wird durch den Einsatz des Statistiklabors gelöst. Ausgewählte Daten werden analysiert. Die entsprechenden Laborseiten werden in den einzelnen Fällen bereitgestellt.
- Report: Es wird ein Musterreport bereitgestellt, der zeigt, wie der Fall mit Hilfe des Statistiklabors hätte gelöst werden können. Der Report wird mit Hilfe der Berichterstellungsfunktion erstellt. Es werden Schritte der Reporterstellung demonstriert.

Die Gestaltung der Startseite wird durch den Charakter der Fallstudie beschrieben. Ist diese fallorientiert, wird lediglich der einzelne Fall auf der Startseite benötigt. Bei einer szenenorientierten Fallstudie spielen andere Aspekte eine Rolle.

<sup>242</sup>Zum Beispiel erfolgt ein Verweis zur Unterscheidung diskreter und stetiger Merkmale.

**Gestaltung einer szenenorientierten Fallstudie** Die Startseite einer szenenorientierten Fallstudie hat das Ziel, dem Studierenden einen Überblick über alle Szenen und Materialien zu geben. Betrachten wir mögliche Gestaltungselemente. Die Beschreibung einzelner Elemente wird durch konkrete Beispiele aus der Fallstudie zur Regressionsanalyse verdeutlicht.<sup>243</sup>

### Bildschirmunterteilung

Um dem Studierenden einen Überblick über die verschiedenen Bereiche zu geben, wird der Bildschirm unterteilt. Auf dem Hauptfeld des Bildschirms erscheint die ausgewählte Szene. In der Kopfzeile wird der Studierende zu einer Einführung in die Problemstellung der Szenenfallstudie geleitet. In der rechten Spalte wird ein Überblick über alle Szenen und ihre Inhalte gegeben. Hyperlinks führen direkt zu den entsprechenden Szenen.

EINFÜHRUNG	SCENEN-NAVIGATOR	HILFE	LEXIKON
<b>FALLSTUDIE REGRESSIONSANALYSE</b>			<b>SCENENÜBERSICHT</b>
<b>DIE STORY</b>			<ul style="list-style-type: none"> <li>▶ <a href="#">Strategische Neuausrichtung</a></li> <li>▶ <a href="#">Von den Daten zum Modell</a></li> </ul>
<p>Das Radsport-Center Garling steht vor schwierigen Entscheidungen: Mit dem Verkauf von Fahrrädern und Zubehör kann Garling kaum noch Gewinn erzielen. Das Unternehmen will sich als Nischenanbieter in Zukunft auf den Verkauf von <a href="#">VeloRain</a> konzentrieren.</p> <p>Mit Hilfe der Regressionsanalyse sollen verschiedene am Markt wirksame Einflussfaktoren untersucht werden und Entscheidungen abgesichert werden.</p>			<p>Fazit: <a href="#">Einführung Lineares Regressionsmodell</a></p> <p>Fazit: <a href="#">Koeffizienten des linearen Regressionsmodells</a></p> <p>Fazit: <a href="#">Lineares Regressionsmodell bei Normalverteilung</a></p>
<b>SCENEN DER FALLSTUDIE</b>			<ul style="list-style-type: none"> <li>▶ <a href="#">Wie gut ist das Modell?</a></li> <li>▶ <a href="#">Das Modell wird erweitert</a></li> <li>▶ <a href="#">Ist das Modell haltbar?</a></li> <li>▶ <a href="#">Die Entscheidung</a></li> </ul>
<p>In 6 Szenen können Sie die Protagonisten bei der Entscheidungsfundung beobachten. Die hier aufgeworfenen Fragen werden in den Lernmodulen aufgegriffen und können von Ihnen im Labor bearbeitet werden.</p> <p>▶ <a href="#">Die Handelnden in der Fallstudie</a></p> <p>▶ <a href="#">Übersicht über alle Szenen: Szenen-Navigator</a></p> <p>Szenen können mit einem Fazit versehen sein. Hier werden die im Lernmodul theoretisch erläuterten Grundlagen auf den Fall angewendet.</p>			<p>Fazit: <a href="#">Beurteilung des linearen Regressionsmodells</a></p> <p>Fazit: <a href="#">Multiples Regressionsmodell</a></p> <p>Fazit: <a href="#">Beurteilung des multiplen Regressionsmodells</a></p>
<b>LERNMODULE</b>			<b>RADSPORT-CENTER GARLING</b>
<p>Zu jeder Szene sind Lernmodule verlinkt, in denen das theoretische Know-How zur Bearbeitung der Fragestellungen vermittelt wird. Links auf die Lernmodule finden Sie unten rechts in der orangefarbenen Box.</p>			<ul style="list-style-type: none"> <li><a href="#">Mitarbeiter</a></li> <li><a href="#">Chronik</a></li> <li><a href="#">Produktpalette</a></li> <li><a href="#">VeloRain</a></li> </ul>
<b>HINWEISE ZUM BEARBEITEN DER FALLSTUDIE</b>			<b>RESSOURCEN</b>
<p>...</p>			<ul style="list-style-type: none"> <li><a href="#">Laborseiten</a></li> </ul>

Abbildung 45: Exemplarische Startseite einer Szene

Zusätzlich wird in der Kopfzeile auf den sogenannten *Szenen-Navigator* verwiesen. Dieser beschreibt detailliert die Ziele der einzelnen Szenen. Es werden die Fragen dargestellt, die im Zusammenhang mit den Szenen aufgeworfen wurden:

<sup>243</sup>Entwicklung der statistischen Inhalte: Projektgruppe Hamburg, Prof. Schlittgen.



Abbildung 46: Szenen-Navigator

Neben den Verlinkungen, die in den vorangehenden Paragraphen angesprochen wurden, werden alle Szenen mit den jeweils relevanten Aufgaben, Beispielen, Theorieteilen, Animationen, Reports der multimedialen Lernsoftware verlinkt. Die Fallstudie sollte so stark wie möglich mit den anderen Komponenten der Lernmodule verknüpft werden, damit der Studierende das in der Multimedia-Software bereitgestellte vielseitige Angebot erkennt und an die Nutzung der unterschiedlichen Komponenten herangeführt wird. Bei dieser starken Verlinkung der einzelnen Komponenten spielt das Guiding eine entscheidende Rolle. Der Studierende benötigt Orientierungshilfen beim Navigieren in dem vernetzten System. Betrachten wir exemplarisch die folgende Szene:



Abbildung 47: Laborverknüpfungen in einer Szenenfallstudie

Durch die dargestellte Verlinkung wird der Studierende direkt auf die zur Verfügung stehende Laborseite geleitet.

### Unabhängigkeit der Komponenten

Jede Szene verweist auf relevante Theoriekomponenten verschiedener Lernmodule. Eine zu enge Verknüpfung der beiden Komponenten ist problematisch, da Studierenden, die sich nur mit einer der beiden Komponenten befassen, Informationen entgehen. Die Theoriekomponente ist unabhängig von der Fallstudie. Zur Veranschaulichung der theoretischen Inhalte können Beispielkomponenten integriert werden. Die Fallstudie kann allerdings nicht unabhängig von der Theorie gestaltet werden, in der Fallstudie kann nicht jede statistische Methode detailliert erklärt werden. Dies würde der motivierenden Rolle der Fallstudie widersprechen, die den Studierenden gerade auf die Theo-



riekomponente hinweisen soll.

Jede Szene der Fallstudie wird mit der entsprechenden Theorieseite verlinkt, die angesprochene Unabhängigkeit gilt nur in der beschriebenen Richtung.

An dieser Stelle wird die Notwendigkeit eines Einsatzkonzeptes deutlich. Es hat die Aufgabe, den Umgang mit einer Fallstudie und die damit verbundenen Ziele festzulegen. Die Einsätze von *Statistik interaktiv komplett* haben gezeigt, dass die Studierenden nicht eigenständig mit den Fallstudien arbeiten. Aus diesem Grund wurde die Fallstudie durch Einsatz des Leitfadens<sup>244</sup> in die Vorlesung integriert. Der Leitfaden enthält alle Aspekte einer Vorlesung, zusammengefasst in einem Dokument. Von diesem wird auf die relevanten Inhalte, Lernmodule, Komponenten verlinkt. Die Bereitstellung dieser Hyperlinks auf die Fallstudien führt dazu, dass die Studierenden diese überhaupt nutzen.

### Vorgehensmodell zur Konzeption und Erstellung einer multimedialen Fallstudie

Das Grundanliegen des Vorgehensmodells besteht darin, die Komplexität des Weges *von A nach B* durch Zerlegung in Einzelschritte zu verringern. Das im Rahmen des Projektes entwickelte Vorgehensmodell zur Konzeption und Erstellung einer Fallstudie setzt sich aus den folgenden Phasen zusammen:

1. Entwurf eines Einsatzkonzeptes, auf dessen Basis die zu entwickelnde Fallstudie eingesetzt werden soll.
2. Suche nach einer geeigneten Story, die alle relevanten Theorieteile abdecken kann und möglichst real ist. Erhebung der benötigten Daten.
3. Einteilung der Fallstudie in einzelne aufeinander folgende Szenen bzw. Ausarbeitung der voneinander unabhängigen Fälle. Zuordnung der statistischen Probleme und Aufgaben der einzelnen Szenen zu der jeweils relevanten theoretischen Komponente.
4. Recherche nach Ressourcen zur Anreicherung der einzelnen Szenen, evt. Selbstdarstellung benötigter Materialien
5. Verknüpfung der Szenen mit den anderen Komponenten anderer Lernmodule (Animationen, Aufgaben, Laborseiten, Beispielen)
6. Konzeption der Startseite in Abhängigkeit von Zielen und Charakter (szenen- versus fallorientiert) der Fallstudie.

Ein Vorgehensmodell für die Gestaltung einzelner Fälle/Szenen zu entwerfen, ist nicht möglich, da diese individuell gestaltet werden müssen.

Nach der Beschreibung der Komponenten der entwickelten Lernsoftware werden im Folgenden die einzelnen Phasen des Entwicklungsprozesses von *Statistik interaktiv komplett* durch ihre Schwerpunkte und Ziele charakterisiert.

## 8.3 Schwerpunkte und Ziele der Phasen des Prototypings

„Wir wollten ein Produkt entwickeln, und wir wollten dieses Produkt gut machen oder ständig verbessern. ... Nicht erst fertigwerden und dann gucken: Ist es gut?“<sup>245</sup>

Die Schwerpunkte und Ziele der phasenweisen Entwicklung von *Statistik interaktiv komplett* werden dargestellt. Dabei wird der Wandel der Funktionalitäten einiger Komponenten diskutiert. Es wird deutlich, dass die Umsetzung verschiedener Entwicklungsideen abhängig von der technischen Realisierbarkeit ist. Es wird davor gewarnt, dass die Technik den Inhalt nicht dominiert.

---

<sup>244</sup>Das Prinzip des Leitfadens wird dem Leser ausführlich im Teil IV erläutert.

<sup>245</sup>[Kindt1999], S. 124.

Der Entwicklungsprozess umfasste drei Phasen, in denen die Prototypen I, II und III<sup>246</sup> entwickelt wurden. Diese werden dem Leser vorgestellt. Zusätzlich werden die Übergänge von einer zur nächsten Phase begründet, um den Leser an der evolutionären Entwicklung teilhaben zu lassen. Es wird deutlich, dass idealtypische Vorgehensweisen neben technischen auch durch politische Variablen bestimmt wurden.

Durch die Entwicklung von Prototypen besteht die Möglichkeit, bestimmte Konzepte und Ideen zu demonstrieren, verschiedene Entwurfsvarianten auszuprobieren und hinsichtlich ihrer Anforderungserfüllung und Akzeptanz zu überprüfen. Während der Entwicklung multimedialer Komponenten einer Lernsoftware ist es zum Beispiel wichtig, zu erproben, welche Gestaltungskonzepte oder Architekturen für den Studierenden zum Erlernen von Statistik geeignet sind. Die Phasen<sup>247</sup> sind in der folgenden Grafik komprimiert dargestellt. Der Entwicklungsprozess des Statistiklabors ist aufgrund seines Umfangs in dieser Darstellung nicht enthalten, sondern wird im Anschluss separat behandelt.

---

<sup>246</sup>Der Prototyp III stellt die endgültige Version von *Statistik interaktiv komplett* dar.

<sup>247</sup>Diese beziehen sich nur auf die in Bielefeld entwickelten multimedialen Lernmodule.

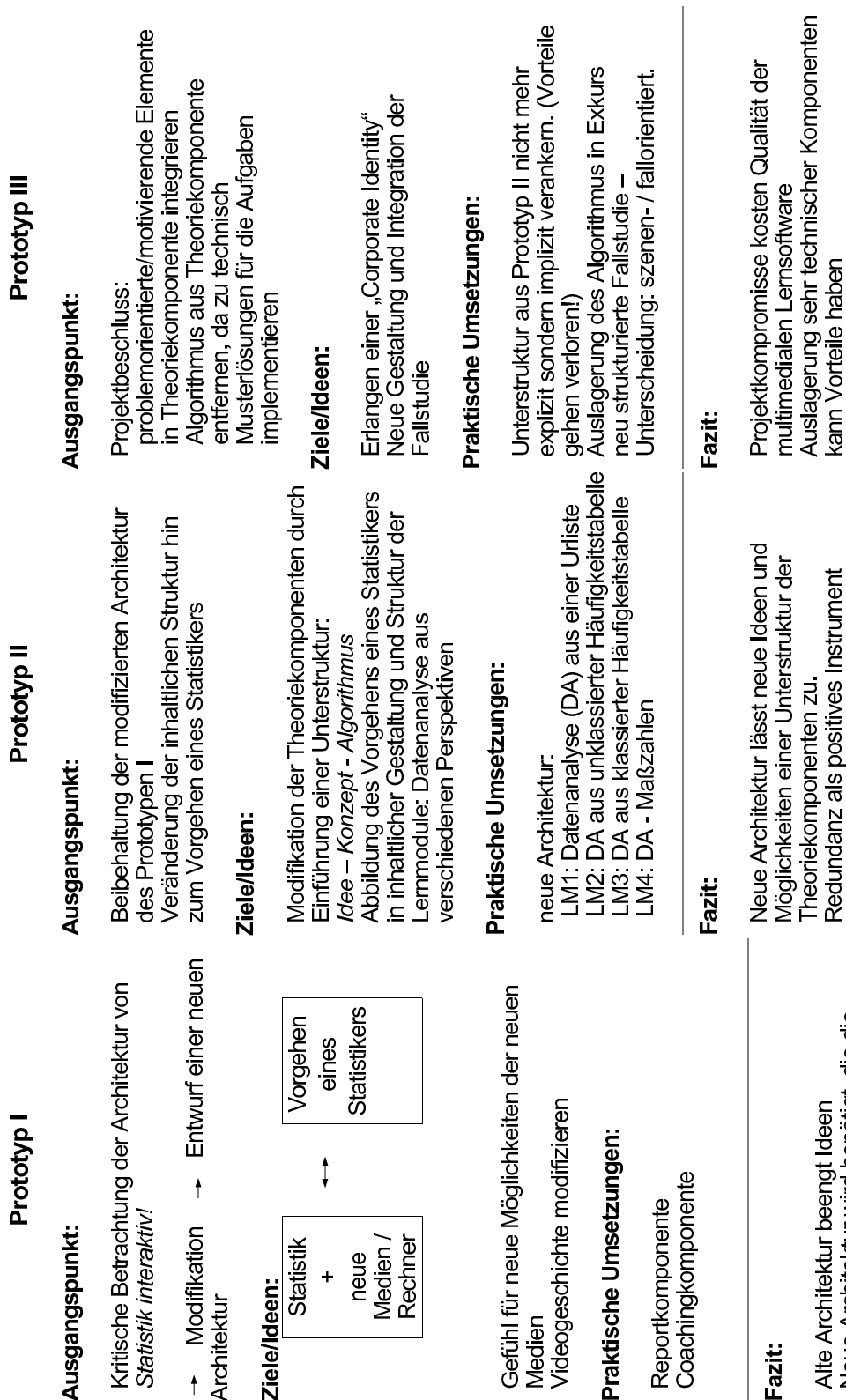


Abbildung 48: Phasen I, II, III des Prototypings

### 8.3.1 Prototyp I - Architektur und Reportsplitter

Der Prototyp I hatte das Ziel, einen Ausschnitt der deskriptiven Statistik so aufzubereiten, dass er die Ideen und Konzepte der Autoren bezüglich der inhaltlichen und strukturellen Gestaltung der neuen Lernmodule transportiert. Der Prototyp hatte zwei sich wechselseitig bedingende Aufgaben, die bei der Gestaltung aufeinander abgestimmt werden mussten:

- Beachtung übergeordneter Ziele, Fragen- und Problemstellungen
- Darstellung detaillierter Gestaltungs- und Strukturideen in übersichtlicher und komprimierter Form

Es wurde der Architekturentwurf für *Statistik interaktiv komplett* entwickelt. Das Resultat, eine modular gestaltete Lernsoftware mit Komponenten verschiedener Funktionalitäten, wurde dem Leser bereits vorgestellt.<sup>248</sup>

Der Prototyp I setzt sich aus den folgenden vier Lernmodulen zusammen: Häufigkeitstabelle, Stabdiagramm, empirische Verteilungsfunktion, Quantile.

Die Pilotprojekte zeigten, dass eine Komponente benötigt wird, die dem Studierenden das Vorgehen eines Statistikers vermittelt. Deshalb wurde mit der Entwicklung einer Komponente *Statistischer Report* begonnen.

Die Schritte der Problemlösung werden in sogenannten *Reportsplittern* dargestellt. Der Prototyp enthält Reportsplitter (kleine Reportteile), die zusammengesetzt den statistischen Report ergeben. Sie bestehen aus Datenanalysen im Statistiklabor und deren Dokumentation. Über die Größe und den thematischen Umfang dieser Splitter wird von Fall zu Fall entschieden. Aufgrund unterschiedlicher Komplexität verschiedener statistischer Methoden erscheint es nicht möglich, hier eine Norm der Art *pro statistischer Methode/pro Lernmodul ein Reportsplitter* vorzugeben. Die Reportsplitter enthalten Verlinkungen zu den entsprechenden Statistiklaborseiten, auf denen die entsprechende Analyse durchgeführt wird.

Im Zusammenhang mit der Komponente *Statistischer Report* wurde eine weitere Komponente eingeführt: die *Coachingkomponente*. Sie leitet den Studierenden. Es existiert z.B. zum Reportverfassen eine Coachingkomponente, die die Idee des Reportverfassens und die Vorgehensweise erklärt. Die Entwicklung und Integration der Coachingkomponente folgt der Idee des Konstruktivismus. Der Studierende soll zwar eigenständig oder *self-paced* lernen, wird aber von dem Dozenten geleitet.

---

<sup>248</sup>Vgl. 8.1.4

► Coaching: Statistisches Reportwriting

*Idee statistisches Reportwriting*

Die Idee des statistischen Reportwritings besteht darin, statistische Methoden gezielt auf ein bestehendes reales Problem anzuwenden. Dabei werden reale Daten beschrieben und das Problem in den statistischen Kontext eingeordnet. Es gilt weiterhin zu erklären, warum bestimmte Methoden in diesem Kontext angewendet werden, warum deren Anwendung überhaupt möglich und auch sinnvoll ist. Die abschließende Aufgabe eines statistischen Reports ist die Formulierung eines Lösungsvorschlags für das ursprüngliche Problem durch die Interpretation der in diesem Zusammenhang bestimmten Analyseergebnisse.

*Schrittweise Entstehung eines statistischen Reports*

1. Ausgangspunkt: Datenmaterial und dazugehöriges Problem/Fragestellung existiert
2. Präzisierung der Fragestellung und Einordnung der Fragestellung in den statistischen Kontext
3. Auswahl geeigneter statistischer Methoden zur Beantwortung der Fragestellung
4. Anwendung der statistischen Methoden auf die Daten beziehungsweise einen Ausschnitt der Daten
5. Kurzdokumentation während des Vorgehens über Gedanken beim Vorgehen
6. Interpretation der durch die Datenanalyse erzielten Ergebnisse
7. Sinnvolle Kombination der Analyseergebnisse und der Interpretation
8. Versuch einer Antwort auf die ausgehende Fragestellung mit möglichst wenigen statistischen Begriffen, d.h. die Antwort soll allgemein verständlich formuliert werden

*Struktur eines statistischen Reports*

1. Problembeschreibung - Fragestellungen
2. Auswahl statistischer Methoden
3. Anwendung der gewählten Methoden auf den vorliegenden Datensatz
4. Interpretation der Datenanalyseergebnisse
5. Rückbezug auf die Realität

*Literaturverweis*

Problem Solving - A statistician's guide, Christopher Chatfield, p.71-74: *Effective report writing*

Abbildung 49: Coachingkomponente

Die Theoriekomponenten des ersten Prototypen enthalten die theoretischen Hintergründe der statistischen Methoden, methodenlogisch strukturiert. Ein Wandel findet in der Entwicklung des folgenden Prototypen statt. Gemäß Phase 3 des Vorgehensmodells findet ein Rückkopplungsprozess statt.

### 8.3.2 Prototyp II - statistische Perspektive

Nach dem Einsatz des Prototypen I wurde deutlich, dass dieser insbesondere in Bezug auf die Gestaltung der Theoriekomponenten verbesserungsbedürftig war. Die Neugestaltung der Theoriekomponenten hatte das Ziel, das statistische Arbeiten mit Daten in den Vordergrund zu stellen, weniger methodenorientiert als im Prototypen I vorzugehen. Dort wurden die Methoden in den vier Lernmodulen lehrbuchartig vermittelt (Häufigkeitstabelle, Stabdiagramm, empirische Verteilungsfunktion, Quantile).

Der Schwerpunkt des Prototypen I, die Kombination von Statistik, statistischem Arbeiten und dem Rechner/den neuen Medien, wurde durch folgende Gestaltungsmaßnahmen weitgeführt:

- Inhaltliche Neugestaltung der Lernmodule: weg von einer methodenorientierten Sicht, hin zu einer statistischen Sicht auf Daten.
- Einführung einer Unterstruktur in jeder Theoriekomponente. Diese wird durch die Einführung eines *Algorithmus* - Abschnittes erweitert.
- Entwicklung einer Fallstudie, die die Vorgehensweise eines Statistikers widerspiegelt.

**Vier neue Lernmodule:** Die inhaltliche Neugestaltung der Lernmodule hat zu der Entstehung von vier neuen Lernmodulen<sup>249</sup> geführt. Diese betrachten die statistische Datenanalyse aus verschiedenen Perspektiven. Die Lernmodule greifen statistische Methoden wiederholt auf.

- Im ersten Lernmodul wird dem Studierenden der Blick eines Statistikers auf Daten in ihrer ursprünglichen Form, d.h. auf die Urliste vermittelt. Es werden ihm statistische Methoden/Instrumente<sup>250</sup> erklärt, mit deren Hilfe er Informationen aus dem nicht aufbereiteten Datenmaterial gewinnen kann - nach dem Prinzip *Look at your data* (Tukey).
- Im zweiten Lernmodul findet eine erste Reduktion der Daten statt. Dem Studierenden wird vermittelt, dass Daten größeren Umfangs mit Hilfe des statistischen Instrumentes *Häufigkeitstabelle* (für diskret skalierte Merkmale) übersichtlich erfasst und dargestellt werden können, dass aber auch Informationen verloren gehen können.<sup>251</sup> Die statistischen Methoden, die im ersten Lernmodul bezüglich der Urliste erklärt wurden, werden im zweiten Lernmodul in Bezug auf die Häufigkeitstabelle vermittelt. Dies ist ein entscheidender Unterschied gegenüber dem Prototypen I. Die statistischen Methoden werden mehrmals aufgegriffen, in verschiedenen Zusammenhängen erklärt, anstatt an einer Stelle (in einem Lernmodul) alle möglichen Varianten einer Methode zu vermitteln. Diese neue Gestaltung zielt darauf ab, die statistischen Methoden so zu vermitteln, wie sie auf bestimmten Reduktionsebenen gebraucht werden. Dass dadurch eine Methode mehrfach in verschiedenen Varianten auftaucht, kann einen positiven Wiedererkennungseffekt beim Studierenden erzielen. Redundanzen werden hier positiv eingeschätzt. Sie sollen den Studierenden bei seiner Wissenskonstruktion unterstützen.
- Analog zur voranstehenden Herangehensweise werden dem Studierenden im dritten Lernmodul die statistischen Methoden in Bezug auf eine klassierte Häufigkeitstabelle vermittelt, d.h. es wird ein weiterer Reduktionsschritt vorgenommen. Er wird darauf hingewiesen, dass die Reduktion Vor- und Nachteile hat. Die Gradwanderung zwischen Informationsverlust und Übersichtlichkeit wird diskutiert.
- Im vierten Lernmodul wird dem Studierenden die Reduktion eines Datensatzes auf eine einzige Zahl vermittelt, statistische Maßzahlen werden eingeführt. Es wird die Nützlichkeit dieser Methode in Bezug auf Vergleiche verschiedener Datensätze hervorgehoben.

In den neu gestalteten Lernmodulen des Prototypen II wird der Studierende von einer Reduktionsebene zur nächsten geführt: von der Datenanalyse des unveränderten Datenmaterials bis hin zur Datenanalyse auf Maßzahlen komprimierter Datensätze. Es werden verschiedene Perspektiven auf die Datenanalyse vermittelt, um bei dem Studierenden das Bewusstsein dafür zu wecken, dass zahlreiche Möglichkeiten existieren, Daten zu analysieren.

**One-page-Principle - Gestaltungszwang durch Technik:** Der Einsatz des ersten Prototypen zeigte, dass die Theoriekomponente nicht übersichtlich genug gestaltet war. Als Konsequenz wurde eine Unterstruktur eingeführt. Die Theoriekomponente setzte sich aus den folgenden Teilen zusammen:

- Idee (Idee der statistischen Methode)
- Konzept (statistische Formalien/Methode)
- Notation (formale Notation)

<sup>249</sup>Vgl. Abb. 48, Phase 2.

<sup>250</sup>Rangwertreihe, Extremwerte, Skalenniveau, Median, Angeln, Fünf-Zahlen-Zusammenfassung, Boxplot, Quartilsabstand, Spannweite, Quantile, arithmetisches Mittel, mittlere quadratische Abweichung, Varianz, Standardabweichung.

<sup>251</sup>z.B. Informationen über die Struktur der Daten. Ein Beispiel: Bei einer ganztägigen Messung des Lärmpegels in einer Straße kann anhand der Urliste erkannt werden, ob und wie der Lärmpegel zu bestimmten Tageszeiten sinkt/steigt. Fasst man die Daten der Messung in einer Häufigkeitstabelle zusammen, geht diese Information verloren, es kann nur noch erkannt werden, wie oft bestimmte Lärmpegel gemessen wurden, aber nicht mehr *wann*.

- Algorithmus (Bereitstellung der R-Befehle zur Umsetzung der entsprechenden statistischen Methode)

Außerdem sollte das statistische Arbeiten im Statistiklabor stärker in den Vordergrund gestellt werden. Jede Methode wurde durch den Algorithmusteil im direkten Zusammenhang mit ihrer technischen Umsetzung vermittelt. Betrachten wir ein Beispiel:

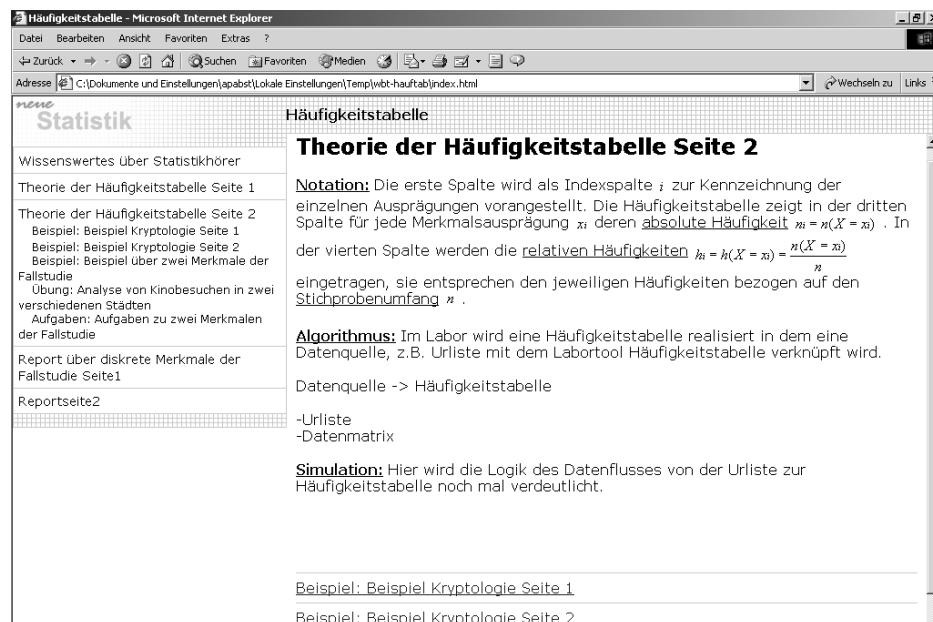
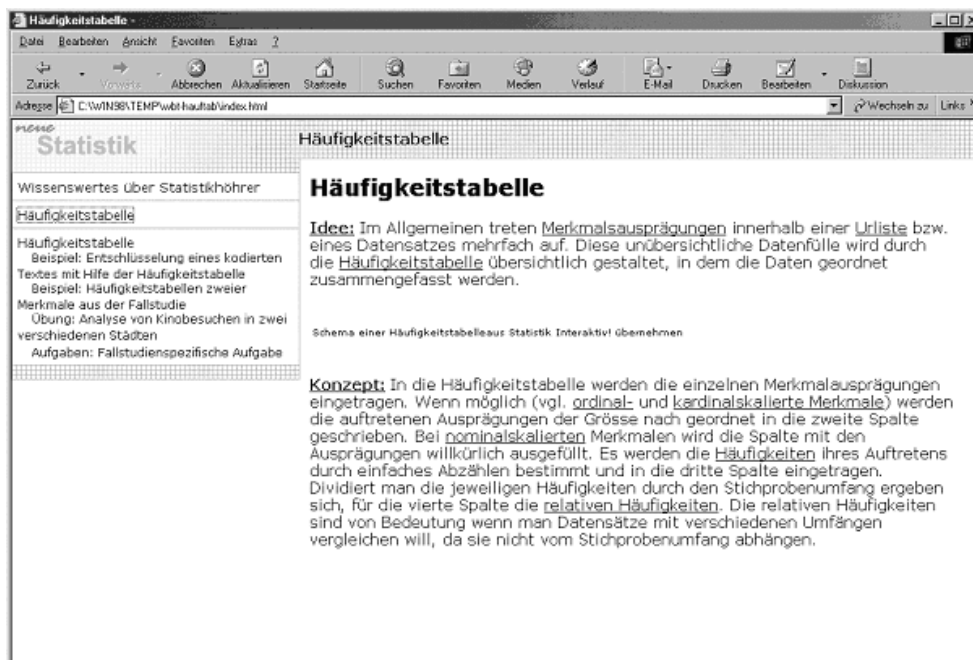


Abbildung 50: Eine Theoriekomponente des zweiten Prototypen



Ein Projektbeschluss<sup>252</sup>, jede Theoriekomponente auf einer Seite abzubilden (One-page-principle), führte zur Auslagerung motivierender Elemente. Dies hatte sogenannte *Motivationskomponenten* zur Folge. Der Vorteil dieser Gestaltung liegt darin, dass die Theoriekomponenten interdisziplinär einsetzbar sind, da sie keine motivierenden Elemente bezüglich einer speziellen Disziplin enthalten. Der Nachteil besteht darin, dass die Inhalte knapp dargestellt werden. Die abgebildete Theoriekomponente erstreckt sich noch über zwei Seiten und ist schon reduziert auf die wichtigsten Aspekte. Auf einer Seite können nur wenige Aspekte der jeweiligen statistischen Methode dargestellt werden. Der spätere Einsatz hat gezeigt, dass die Theoriekomponenten durch andere Materialien erweitert werden müssen, um die theoretischen Inhalte zu vermitteln. Dies stellt eine Schwäche der Lernsoftware dar.

Während der Entwicklung des Prototypen II wurde deutlich, dass es problematisch ist, die einzelnen Lernmodule so zu entwickeln, dass sie unabhängig voneinander einsetzbar sind. Das ursprüngliche Ziel bestand darin, die Lernmodule so zu gestalten, dass sie in beliebiger Reihenfolge vom Studierenden im Lernprozess eingesetzt werden können. Dies darf nicht auf Kosten logischer Reihenfolgen geschehen. Gerade bei umfangreichen Lernmodulen wie *Regression* ist dieses Ziel fragwürdig, da viele Grundkenntnisse benötigt werden. Eine bestimmte Reihung in der Betrachtung der Lernmodule scheint daher unumgänglich.

### 8.3.3 Prototyp III - konträre Prinzipien

Die im Folgenden beschriebenen Veränderungen sind zum Großteil eine Folge projektinterner Diskussionen. Die Theoriekomponente ist stark betroffen. Motivierende Elemente werden wieder aufgenommen, die Unterstruktur entfernt und der Algorithmusteil in einen Exkurs ausgelagert. Dies geschieht im Zusammenhang mit übergeordneten Prinzipien.

Während des Versuches, die Qualität der multimedialen Lernsoftware schrittweise im Entwicklungsprozess zu verbessern, müssen die Zielsetzung und grundsätzliche Prinzipien einbezogen werden. Das Problem bestand darin, dass man sich zu Beginn des Forschungsprojektes *Neue Statistik* nicht auf grundsätzliche Prinzipien geeinigt hatte. Der Übergang vom Prototypen II zum Prototypen III zeigt, welche Fragen im Entwicklungsprozess aufgetreten sind, die je nach übergeordneter Zielsetzung verschieden gelöst werden können. Die wesentlichen Aspekte betreffen die drei Prinzipien *Corporate Identity*, *Technische Reporterstellung* und *Nachhaltigkeit*, die im Folgenden schrittweise diskutiert werden.

**Corporate Identity:** Im Projekt *Neue Statistik* wird darunter die Notwendigkeit verstanden, ein einheitliches Produkt herzustellen. Die multimedialen Lernmodule von *Statistik interaktiv komplett*, die von unterschiedlichen Projektpartnern erstellt werden, sollen sich zu einem einheitlichen, stimmigen Ganzen zusammenfügen. Dies betrifft die äußere Form und die inhaltliche Gestaltung. Mit der Architektur von *Statistik interaktiv komplett* wird das Ziel verfolgt, durch die Lernmodule eine einheitliche Vermittlung der Statistik als Ganzes durchführen zu können.

Die Festlegung der Komponenten bestimmt den Inhalt. So ist die Entscheidung darüber, ob die Motivation der Studierenden aus den Theoriekomponenten in eine Fallstudie ausgelagert werden soll, entscheidend für die Gestaltung der Texte der Theoriekomponente. Enthält die Theoriekomponente motivierende Elemente, legt sie sich auf eine Disziplin fest. Lagert man sie aus, ist sie breiter einsetzbar, motiviert aber nicht mehr zum Erlernen der Theorie. Diese Aufgabe übernehmen dann andere Komponenten (Fallstudien oder Beispiele, Aufgaben). Um eine Corporate Identity zu erlangen, ist es entscheidend, dass diese Aspekte in allen Lernmodulen und ihren Komponenten einheitlich umgesetzt werden.

**Inkompatibilität übergeordneter Prinzipien:** Im Entwicklungsprozess hat sich herausgestellt, dass bestimmte Prinzipien schwer kompatibel sind: das One-page-Principle, das gegen das Scrollen auf Bildschirmseiten der Lernsoftware spricht, ist schwer vereinbar mit der Integration motivieren-

<sup>252</sup>Eine Begründung ist dem Autor nicht bekannt.

der Elemente in die Theoriekomponenten. Auch das Prinzip der interdisziplinären Verwendbarkeit der Theoriekomponenten steht hierzu im Widerspruch. Bei dem Versuch, eine Corporate Identity zu erlangen, muss also vorher festgelegt werden, welche Prinzipien Priorität bei der Entwicklung haben. Ansonsten führt die Entwicklung trotz festgelegter Normen nicht zu einer Corporate Identity. Im Prototypen III werden das One-page-Principle und das Prinzip der interdisziplinären Verwendbarkeit der Theoriekomponenten verletzt, da Motivationsteile in die Theoriekomponenten integriert wurden.

**Technische Reporterstellung:** Es gibt verschiedene Möglichkeiten, statistische Reports mit Hilfe der Lernsoftware zu erstellen. Die erste besteht darin, die Komponenten Statistiklabor und Editor kombiniert einzusetzen, im Statistiklabor die statistischen Analysen durchzuführen, diese in den Editor zu übertragen und dort zu interpretieren. Die zweite Möglichkeit besteht darin, die gesamte Reporterstellung im Statistiklabor durchzuführen. Hierzu wurde die Berichterstellungsfunktion<sup>253</sup> entwickelt und im Prototypen III implementiert. Diese übernimmt die Aufgabe, innerhalb des Labors ein Dokument zu erstellen, welches ausgewählte Teile der statistischen Datenanalyse aus dem Labor in einen Bericht überträgt. In diesem besteht die Möglichkeit, Interpretationen und Bemerkungen hinzuzufügen. Die Berichterstellungsfunktion wurde aufgrund von Kritik der Studierenden bezüglich des kombinierten Einsatzes zweier Komponenten (Statistiklabor und Editor) entwickelt und in das Statistiklabor integriert. Der alte Editor wurde beibehalten. Er ermöglicht die direkte Dokumentation auf einer Laborseite. Bei einigen Lösungsvorschlägen in Teil I<sup>254</sup> wurde im Texteditor z.B. der Umgang mit den im R-Kalkulator bereitgestellten Funktionen erklärt. Betrachten wir im Folgenden eine Statistiklaborseite, aus der mit Hilfe der Berichterstellungsfunktion ein Report erstellt werden kann:

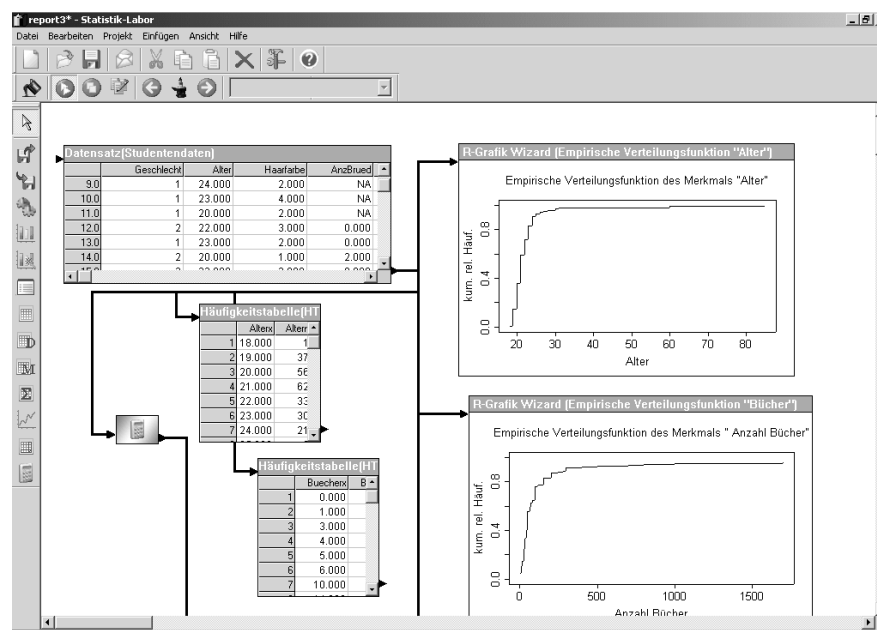


Abbildung 51: Statistiklaborseite als Basis eines Reports

In dem unten abgebildeten Wizard der Berichterstellung können Objekte der Laborseite ausgewählt und somit in den Report übernommen werden. Es öffnet sich ein Dokument, welches die ausgewählten Laborobjekte enthält. In diesem kann die Dokumentation stattfinden.

<sup>253</sup>Vgl. Abb. 52.

<sup>254</sup>Vgl. z.B. Abb. 23.

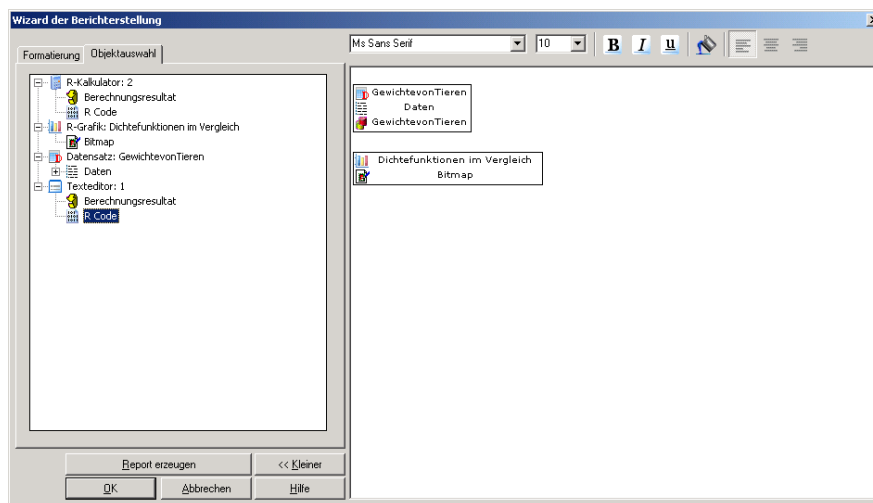


Abbildung 52: Reporterstellung mit Hilfe der Berichterstellungsfunktion

**Nachhaltigkeit und Modifizierbarkeit:** Die Nachhaltigkeit der multimedialen Lernsoftware *Statistik interaktiv komplett* stellt eines der übergeordneten Ziele des Projektes *Neue Statistik* dar.<sup>255</sup> Nachhaltigkeit stellt an die zu entwickelnde multimediale Lernsoftware den Anspruch, lange über ihren Erstellungszeitraum hinaus einsetzbar zu sein. Um eine breite Masse anzusprechen, muss die Lernsoftware so entwickelt werden, dass sie nicht nur in den Wirtschaftswissenschaften, sondern auch in die Curricula der Soziologie oder Medizin integriert werden kann. Dies betrifft hauptsächlich die Auslagerung der Motivationselemente. Für diesen Zweck ist es notwendig, die Lernmodule modifizierbar zu gestalten. Inhalte müssen austauschbar sein. Nachhaltigkeit kann zudem bedeuten, dass der Umgang mit der Lernsoftware einfach ist. Die multimedialen Lernmodule müssen problemlos ohne großen Aufwand individuelle Veränderungen durch den jeweiligen Dozenten zulassen.<sup>256</sup>

### Umsetzungen im Prototypen III

Der Prototyp III wurde mit dem Ziel der Erreichung einer Corporate Identity entwickelt. Größtenteils ist die Theoriekomponente von den Weiterentwicklungen betroffen.

Die übersichtliche Strukturierung der Theoriekomponente in *Idee*, *Konzept*, *Notation* und *Algorithmus* tritt nur noch implizit auf. Außerdem wurden motivierende Elemente wieder integriert, da erkannt wurde, dass dies zu einer nicht tragbaren Reduktion der Inhalte führte. Das *One-Page-Principle* wurde nicht weiter verfolgt. Die motivierenden Elemente verlängern die Theoriekomponenten wieder. Auch die interdisziplinäre Einsatzfähigkeit ist nicht mehr ausnahmslos gegeben. Die Übersicht und die Wiedererkennungseffekte, die durch das explizite Vorhandensein der Strukturierung geschaffen wurden, sind verloren gegangen. Begründet wurden diese sogenannten Verbesserungen oder Weiterentwicklungen damit, dass es einen Studierenden langweilt, in der Theoriekomponente auf die reine Theorie zu treffen (deshalb die Integration motivierender Elemente).

Der Einsatz zeigte, dass die Theoriekomponenten weiterhin zu knapp gestaltet waren und vom Dozenten deshalb durch ein Skript mit theoretischen Ausführungen ergänzt werden mussten. Den Studierenden fehlte zudem eine einheitliche Struktur der Theoriekomponenten. Die Abschaffung der Struktur des Prototypen II kann als Fehlentscheidung bezeichnet werden. Zudem verhindern die motivierenden Elemente den interdisziplinären Einsatz der Theoriekomponenten.

Der Prototyp II motivierte über die Fallstudie. Im Prototypen III aber verschiebt sich das Gewicht

<sup>255</sup> Vgl. 11

<sup>256</sup> Nicht nur Lernen, sondern auch Lehren ist individuell.

hin zur Theoriekomponente. Diese enthält auch motivierende Elemente, was zur Folge hat, dass die Fallstudie in ihrer Bedeutung in den Hintergrund tritt.

Die Entwicklung und Implementierung der Berichterstellungsfunktion im Statistiklabor kann als positive Weiterentwicklung bewertet werden. Für die Studierenden bedeutet dies eine technisch einfache Erstellung ihrer Reports. Um sie mit der Form der Darstellung vertraut zu machen, wurden auch alle implementierten multimedialen Reports mit Hilfe der Berichterstellungsfunktion erstellt.

#### 8.4 Beschreibung einzelner Entwicklungsschritte des Statistiklabors

Das Statistiklabor ist die zentrale Komponente der Lernsoftware *Statistik interaktiv komplett*. Sie wurde kontinuierlich weiterentwickelt. Die schrittweise Verbesserung der Funktionalitäten des Statistiklabors lässt sich nicht streng den drei beschriebenen Phasen zuordnen. In seiner Endversion, in der es den Mediaprix 2003<sup>257</sup> gewonnen hat, wurde das Statistiklabor wie folgt beschrieben:<sup>258</sup>

Das Statistiklabor ist eine explorative und flexibel anzupassende Lern- und Arbeitsumgebung für die Statistik-Grundausbildung unterschiedlicher Studiengänge. Als didaktisch motiviertes Produkt unterstützt das Statistiklabor eine qualitative Veränderung der Lehre durch folgende Prinzipien:

- Sicherstellung eines problem- und fallorientierten Zugangs zu statischen Methoden und Konzepten
- Unterstützung explorativen, selbstständigen und fehlertoleranten Lernens auf der Basis frei disponibler oder auch vorgegebener Problemszenarien
- Angeleitetes Lernen durch die Präsentation und Bearbeitung von Aufgaben und Musterlösungen
- Hohe Adaptivität an den individuellen Lernprozess
- Flexible Integration in unterschiedliche Einsatzkonzepte

Während in der konventionellen Lehre häufig eine formale Vermittlungsperspektive eingenommen wird, unterstützt das Statistiklabor ein problemorientiertes Arbeiten an realen Fragestellungen. Studierenden wird das vollständige Bearbeiten statistischer Fragestellungen in einer Umgebung ermöglicht und ihnen dadurch der Zugang zum Fach Statistik erleichtert. Hochschullehrern gibt das Statistiklabor einen umfangreichen Werkzeugkasten in die Hand, der flexibel an die eigenen Lehrinhalte und -strategien angepasst werden kann. Das Statistiklabor kann ergänzend zur Präsenzlehre eingesetzt und auch vollständig in netzgestützte Lernszenarien eingebunden werden.

Die Beschreibung verdeutlicht, dass sowohl die Studierenden als auch die Dozenten als Benutzer betrachtet werden. Aus diesem Grund wurden im Entwicklungsprozess die Bedürfnisse beider berücksichtigt. Das Statistiklabor ist konzipiert worden, um die Grundausbildung in Statistik substantiell zu verbessern. Bei der Entwicklung wurde vor allem auf folgende Punkte Wert gelegt:<sup>259</sup>

- Skalierbarkeit: Sowohl einfache als auch komplexe statistische Szenarien lassen sich im Statistiklabor bearbeiten. So eröffnen sich vielfältige Einsatzmöglichkeiten von einer Einführung in die deskriptive Statistik über die Simulation theoretischer Probleme aus dem Bereich Schätzen oder der Regressionsanalyse bis hin zur Vermittlung der statistischen Programmierung in R.
- Didaktisches Fundament: Authentische statistische Probleme können durch die Integration realer Daten leicht im Statistiklabor bearbeitet werden. Der offene Aufbau des Arbeitsplatzes unterstützt dabei individuelle Herangehensweisen. Nicht die Berechnung und Anwendung

---

<sup>257</sup> <http://www.mediaprix.org/>

<sup>258</sup> Vgl. [www.statistiklabor.de](http://www.statistiklabor.de). Diese Beschreibung ist abstrakt und wird im Anschluss anhand von Beispielen konkretisiert.

<sup>259</sup> Vgl. Webauftritt des Statistiklabors [www.statistiklabor.de](http://www.statistiklabor.de)

formalen Wissens steht im Vordergrund, sondern die Erarbeitung konstruktiver Lösungsstrategien.

- **Fachdidaktisches Fundament:** Wesentlicher Ansatzpunkt für eine Verbesserung der Statistiklehre ist die im Statistiklabor verwirklichte Datenorientierung statt einer einseitigen Methodenorientierung. Interaktive statistische Experimente und Simulationen erleichtern das Verständnis theoretischer Probleme und können leicht vom Hochschullehrer selbst erstellt und an die eigenen Bedürfnisse angepasst werden.
- **Flexibilität im Einsatz:** Das Statistiklabor ist in der gesamten Grundausbildung flexibel einsetzbar. Es kann zur visuellen Unterstützung in der Vorlesung, als Werkzeug in Tutorien und Übungen, zum Selbstlernen zu Hause oder auch als Prüfungsumgebung eingesetzt werden.
- **Akzeptanz und Nachhaltigkeit:** Die Integration der verbreiteten Programmiersprache R ermöglicht die Nutzung des Statistiklabors ohne erneuten Einarbeitungsaufwand. Das Expertenwissen einer weltweiten Entwicklergemeinschaft steht hier zur Verfügung. Die Funktionen des Statistiklabors (der statistische Arbeitsplatz) werden bereits seit 1997 konsequent weiterentwickelt und verbessert.

### Funktionen des Statistiklabors im Überblick:

#### Der statistische Arbeitsplatz

Nach dem Start des Statistiklabors wird ein Arbeitsblatt geöffnet, auf dem statistische Objekte wie Datensätze, Häufigkeits- und Kontingenztabelle oder Zeitreihen per Mausklick platziert werden können. Über einen Generator können Zufallszahlen erzeugt, über einen Datenimport reale Datensätze im Excel- oder ASCII-Format importiert werden.

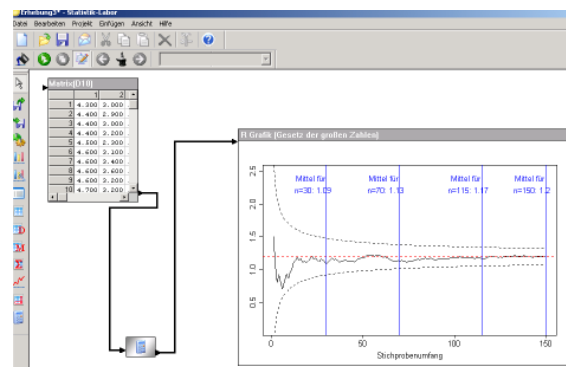


Abbildung 53: Statistischer Arbeitsplatz des Labors

#### Laborobjekte und Datenfluss

Betrachten wir die Objekte des Statistiklabors im Überblick:

- Datenquellen und Datenpräsentation zur Erzeugung, Darstellung bzw. zum Import von Daten (Datensatzimport, Zufallszahlen-Generator, Urliste, Datensatz, Matrix)
- Datenanalyse zur Analyse bzw. Manipulation von Daten (Häufigkeitstabelle, Kontingenztabelle, Zeitreihen, R-Kalkulator)
- Datenausgaben zur grafischen oder textuellen Ausgabe von Daten und zum Export in Fremdformate (Datensatzexport, R-Grafik, R-Grafik-Wizard, Texteditor)

Mit Hilfe von Konnektoren (schwarze Pfeile in Abb. 53) können diese verschiedenen Objekte miteinander verbunden werden und die entsprechenden Datenflüsse sowohl logisch festgelegt als auch grafisch dargestellt werden. Daten aus einem Objekt können so zur weiteren Analyse in ein neues

Objekt übergeben werden. Damit wird ein schrittweises Bearbeiten umfangreicher Fragestellungen unterstützt. Die Logik und die Abfolge des Datenflusses können vom Studierenden oder vom Dozenten je nach Aufgabenstellung definiert werden. Im Einsatz erkannten wir, dass die Datenflüsse im Labor dem des statistischen Arbeitens entsprechen müssen. Die Studierenden hatten Schwierigkeiten, nachzuvollziehen warum die Grafik, in der ein Stabdiagramm erzeugt wurde, durch einen Konnektor mit dem Objekt *Datensatz* verbunden war. Das statistische Arbeiten verlangt einen Zwischenschritt: die Erstellung einer Häufigkeitstabelle. Als Konsequenz wurden die Datenflüsse und Laborobjekte im Labor an die Schritte des statistischen Arbeitens angegliedert.

### Einfache Visualisierungen

Mit Hilfe des Grafik-Wizards können Standard-Diagramme wie Histogramme, Boxplots oder Streudiagramme erzeugt werden.

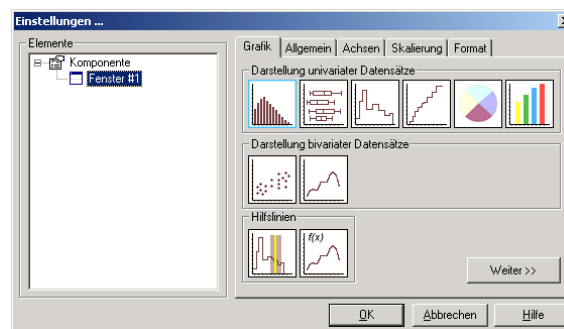


Abbildung 54: Visualisierungsmöglichkeiten des Grafik-Wizards

Der Benutzer kann zwischen den Möglichkeiten zur grafischen Darstellung univariater und bivariater Datensätze auswählen. Es werden keine R-Befehle zur Erstellung der Grafiken benötigt. Im Einsatz stellte sich heraus, dass dies insbesondere dazu geeignet ist, die Studierenden an das Arbeiten im Statistiklabor heranzuführen.

### Integration von R

Die statistische Programmiersprache R ist vollständig im Statistiklabor verfügbar und wird über den R-Kalkulator aufgerufen. Diese Schnittstelle bietet vor allem Statistikexperten flexible Möglichkeiten zur Bearbeitung statistischer Probleme. Simulationen und interaktive Experimente, die normalerweise als Java-Applets aufwendig programmiert werden müssen, können mit Hilfe von R durch Hochschullehrer selbst erstellt werden.

### Statistiktaschenrechner

Ein Statistiktaschenrechner ermöglicht Anfängern, ohne Grundkenntnisse der Syntax einfache Anweisungen im R-Code zu erstellen. Dies kann vor allem zur Einführung in die Programmierung mit R genutzt werden oder auch zum Bearbeiten einfacherer Fragestellungen.

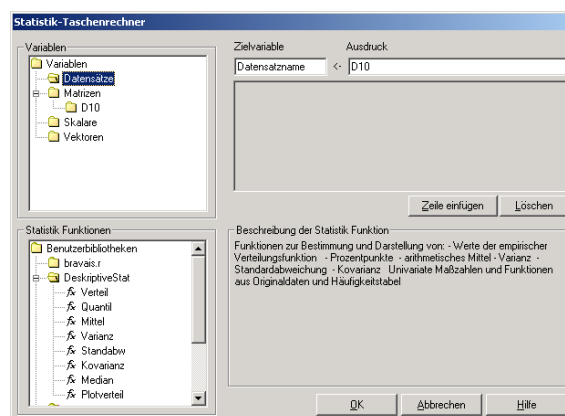


Abbildung 55: Der Statistiktaschenrechner

Als *Zielvariable* kann für die Analyse ein Datensatz ausgewählt und bezeichnet werden. Die Auswahl führt dazu, dass im R-Kalkulator auf diesen Datensatz zugegriffen werden kann. Der Taschenrechner stellt zudem (in dem linken unteren Bereich der Grafik) sogenannte *Benutzerbibliotheken* bereit. Dozenten können über Benutzerbibliotheken individuelle Szenarien abbilden und die Funktionalität des Statistiklabors selbst erweitern. Anweisungen und Berechnungen sind für Studierende über Benutzerbibliotheken in kompakter Form ohne Grundkenntnisse von R nutzbar. Die Bibliotheken werden im R-Kalkulator aufgerufen: der R-Code bleibt dem Studierenden verborgen. Diese Gestaltung wird als *Schichtenprinzip* bezeichnet. Informationen werden durch die Einführung verschiedener Schichten vor dem Studierenden versteckt. Im R-Kalkulator sieht er lediglich den Funktionsaufruf. Die Bibliothek erst enthält den R-Code dieser Funktion. Die im Teil I präsentierten Lösungsvorschläge wurden nach diesem Prinzip gestaltet.<sup>260</sup>

Im Einsatz wurden verschiedene Bibliotheken benötigt. Jeder Projektpartner stellte entwickelte Bibliotheken in der Webpräsenz zur Nutzung bereit. Dieses Forum an Bibliotheken wird sich im Laufe des Einsatzes des Statistiklabors kontinuierlich vergrößern, wenn verschiedene Dozenten ihre erstellten Bibliotheken bereitstellen.

### Kommunikation im Entwicklungsprozess

Die schrittweise Verbesserung des Statistiklabors wurde zentral organisiert. Die an der Laborentwicklung beteiligten Projektgruppen setzten das Labor in der Lehre ein und meldeten Fehlerquellen an CeDiS. Jeder identifizierte Fehler wurde nummeriert und in die Webpräsenz<sup>261</sup> gestellt, um von den anderen Projektpartnern überprüft werden zu können. Die anschließende Behebung wurde in der Webpräsenz kommuniziert. Zudem wurden auf diesem Weg Ideen und Wünsche zur Laborgestaltung kommuniziert.

Während des Entwicklungsprozesses entstanden schrittweise verschiedene Versionen des Statistiklabors. Diese Versionen wurden zum Download in der Webpräsenz bereitgestellt. Es wurde den beteiligten Projektpartnern eine Zeitspanne zum sogenannten *Betatesting* eingeräumt. Die identifizierten Fehler wurden diskutiert und anschließend verbessert. Auf diese Weise entstanden schrittweise verbesserte Versionen des Statistiklabors.

<sup>260</sup> Vgl. z.B. 4.3.

<sup>261</sup> Es handelt sich um die Webpage des Projektes *Neue Statistik*, die als Kommunikationsgrundlage der Projektpartner während des Entwicklungs- und Einsatzprozesses von *Statistik interaktiv komplett* diente. <http://www.neuestatistik.de/>

**Meilensteine im Entwicklungsprozess des Statistiklabors** Es können folgende Meilensteine<sup>262</sup> im Entwicklungsprozess des Statistiklabors identifiziert werden:

- Übernahme des Labors aus *Statistik interaktiv!*  
Das Labor der Vorgängersoftware wurde zunächst übernommen. Diese hatte allerdings erhebliche Mängel (z.B. keine stabile Lauffähigkeit), die schrittweise behoben wurden.
- Entwicklung einer benutzerfreundlichen Oberfläche  
Die Benutzeroberfläche wurde schrittweise an die den Benutzern bekannten Windowsstandards angeglichen.
- Bereitstellung und Dokumentation von Laborseiten  
Es erfolgte die Erkenntnis, durch die Bereitstellung dokumentierter Laborseiten Transparenz und Nachvollziehbarkeit gewährleisten zu müssen. Als Konsequenz wurde jede bereitgestellte Laborseite dokumentiert.  
Es folgt ein Beispiel einer dokumentierten Statistiklaborseite:<sup>263</sup>

---

<sup>262</sup>Der Begriff Meilenstein wird hier in zwei Definitionen verwendet: *wichtiger Bestandteil* und *etwas zu einem Zeitpunkt zu Erreichendes*.

<sup>263</sup>Diese wurde in einer Zwischenversion des Labors erstellt, enthält daher die Bezeichnung *LDL-Kalkulator*, die später durch *R-Kalkulator* abgelöst wurde.



Dokumentation zur Laborseite:

**Inhalt:** Mehrmaliges Werfen einer selbst zu bestimmenden Anzahl von Münzen. Es wird gezählt, wie oft jeweils „Kopf“ gefallen ist. Vergleich zwischen Ergebnis des Experiments und Theorie.

**Lernziel:** Wie verhalten sich Experiment und theoretische Verteilung zueinander?

**Vorlesung 7.01. Statistik-Motivation (Teil b)**

Anteilvektor	n	h	F
1	3	0,25	0,06
2	1	0,25	0,19
3	2	0,25	0,44
4	3	0,25	0,26
5	4	0,25	0,06
6			

**Interaktive Komponenten:**

- (1) LDL-Kalkulator (Parameter): Der Parameter  $n$  dient dazu, anzugeben, wie viele Würfe durchgeführt werden sollen. Mit dem Parameter  $k$  kann man beeinflussen, wie viele Münzen in jedem der  $n$  Durchgänge geworfen werden.

**Sonstige Komponenten:**

- (2) LDL-Kalkulator (Münzwürfe): Durchführung der Serie von Münzwürfen. In der Variablen *Anteilvektor* wird festgehalten, wie oft in jedem Durchgang „Kopf“ gefallen ist.
- (3) Häufigkeitstabelle (Anteilvektor): Stellt die Häufigkeitstabelle der erzielten Ergebnisse dar.
- (4) Koordinatennetz (Stabdiagramm): Enthält das Stabdiagramm der erzielten Ergebnisse.
- (5) Koordinatennetz (empirische Verteilungsfunktion): Enthält die Verteilungsfunktion der erzielten Ergebnisse.
- (6) LDL-Kalkulator (theoretische Wahrscheinlichkeiten): Die theoretischen Wahrscheinlichkeiten für das Eintreten der einzelnen Ergebnisse werden in *theor* gespeichert.
- (7) LDL-Kalkulator (Vorber. Stabd./Verteilungsf.): Technische Zwischenschritte, um die theoretischen Wahrscheinlichkeiten in Stabdiagramm und Verteilungsfunktion darstellen zu können.
- (8) Koordinatennetz (Stabdiagramm (Theorie)): Darstellung des theoretischen Stabdiagramms.
- (9) Koordinatennetz (theoretische Verteilungsfunktion): Darstellung der theoretischen Verteilungsfunktion.

Abbildung 56: Dokumentation einer Statistikkolaborseite

Anhand der Dokumentation erkennt der Benutzer die Ziele der verschiedenen Laborobjekte. Es wird ersichtlich, welche Objekte die Möglichkeit der Interaktion bieten.

- Implementierung der Berichterstellungsfunktion  
Die Berichterstellungsfunktion ist wie oben beschrieben aus dem Editor hervorgegangen. Sie ermöglicht die Kombination statistischer Analysen und Interpretationen zu einem statistischen Report. Der Einsatz hat gezeigt, dass die Implementierung der Berichterstellungsfunktion notwendig war. Die Studierenden erstellten bereitwillig statistische Reports. Zuvor war dies mit technischen Problemen verbunden und schreckte ab. Ein Vorteil der Berichterstellungsfunktion ist ihr Aufruf im Labor. Die Datenanalyse wird auch im Labor durchgeführt. Die Reporterstellung kann also innerhalb einer Komponente durchgeführt werden.

- Abbildung des Arbeitsablaufes eines Statistikers durch Einsatzmöglichkeiten und Nutzung der Objekte in einer bestimmten Reihenfolge. Dem Leser wurde bereits ein Beispiel gegeben: *Im Einsatz erkannten wir, dass die Datenflüsse im Labor dem des statistischen Arbeitens entsprechen müssen. Die Studierenden hatten Schwierigkeiten nachzuvollziehen, warum die Grafik, in der ein Stabdiagramm erzeugt wurde, durch einen Konnektor mit dem Objekt Datensatz verbunden war. Das statistische Arbeiten verlangt einen Zwischenschritt: die Erstellung einer Häufigkeitstabelle. Als Konsequenz wurden die Datenflüsse und Laborobjekte im Labor an die Schritte des statistischen Arbeitens angeglichen.*
- Stabilisierung der Lauffähigkeit des Statistiklabors:  
Das Statistiklabor lief weder in *Statistik interaktiv!*, noch in den ersten beiden Prototypen stabil. Die in dieser Arbeit bereitgestellte Version läuft stabil.
- Einbindung von Bibliotheken und Implementierung der Exkurs-Komponente:  
Die Möglichkeit, im Statistiklabor eigene R-Funktionen in Form von Bibliotheken einzubinden, erweiterte die Funktionalitäten und die Einsatzmöglichkeiten des Labors in hohem Maße. Dozenten können die Laborseiten durch die Einbindung von Bibliotheken nach ihren individuellen Bedürfnissen gestalten.

**Ausgewählte Detailentwicklungen des Statistiklabors** Im Anschluss werden dem Leser ausgewählte Detailentwicklungen des Statistiklabors vorgestellt. Diese betreffen die folgenden Aspekte:

- Reihenfolge und Funktionalitäten der Objekte
- Der Exkurs *Datenanalyse.R*
- Erstellung statistischer Reports
- Der Texteditor als Laborobjekt
- Funktionalitäten des Grafik-Wizards
- Das Laborobjekt Urliste
- Das Laborobjekt Kontingenztafel

**Reihenfolge und Funktionalitäten der Objekte:** Betrachten wir exemplarisch ein Problem. Es war in der im März 2003 aktuellen Version nicht möglich, den Grafik-Wizard direkt mit der Häufigkeitstabelle zu verbinden. Der Grafik-Wizard stellt verschiedene Möglichkeiten bereit, statistische Grafiken zu erstellen, z.B. das Histogramm. Das Histogramm konnte nur dann erstellt werden, wenn die Datenquelle durch einen Konnektor mit dem Grafik-Wizard verbunden wurde. Dies entspricht nicht dem statistischen Arbeiten. Der Statistiker komprimiert die Daten in der klassierten Häufigkeitstabelle und entwickelt aus dieser (nicht direkt aus der Datenquelle) das Histogramm. Im sich anschließenden Entwicklungsschritt wurde diese Schwäche des Labors behoben. In der Endversion wird das Histogramm durch die Verbindung der Häufigkeitstabelle mit dem Grafik-Wizard erstellt. Das Beispiel zeigt eine Schwäche des Entwicklungsprozesses: Technik und Statistik wurden getrennt betrachtet.

In diesem Zusammenhang entstand ein weiteres Problem. Der Grafik-Wizard stellt die Möglichkeit bereit, eine Häufigkeitstabelle zu erzeugen. Im Einsatz stellte sich heraus, dass die Klassenbreiten beim Erstellen einer klassierten Häufigkeitstabelle nicht variiert werden konnten, d.h. es konnte nur eine Klassenbreite für die gesamte Häufigkeitstabelle angegeben werden. Außerdem konnte der Benutzer keinen Startwert für die Klasseneinteilung festsetzen. Während des Einsatzes wurde deshalb statt des Grafik-Wizards der R-Kalkulator in Verbindung mit dem Objekt R-Grafik genutzt. Durch die Entwicklung der Bibliothek *Datenanalyse.R* bestand nun die Möglichkeit, die Klassenbreiten zu variieren und einen Startpunkt zu bestimmen. Als Konsequenz des Einsatzes wurde der Grafik-Wizard überarbeitet. Zudem wurde *Datenanalyse.R* in Form eines Exkurses implementiert.

Der Exkurs *Datenanalyse.R*<sup>264</sup> Das Statistiklabor wurde eingesetzt, um statistische Aufgaben zu lösen. Es stellte sich heraus, dass die Studierenden Hilfen zum Umgang mit den Laborobjekten benötigten. Dies betraf insbesondere den R-Kalkulator, die Eingabe von R-Befehlen. Als Konsequenz wurde während des Einsatzes ein Paper erstellt, in dem den Studierenden der Umgang mit jedem neu genutzten Laborobjekt und jeder neue R-Befehl erläutert wurde. In der sich anschließenden Entwicklungsphase wurde *Datenanalyse.R* als Exkurs wie folgt implementiert:

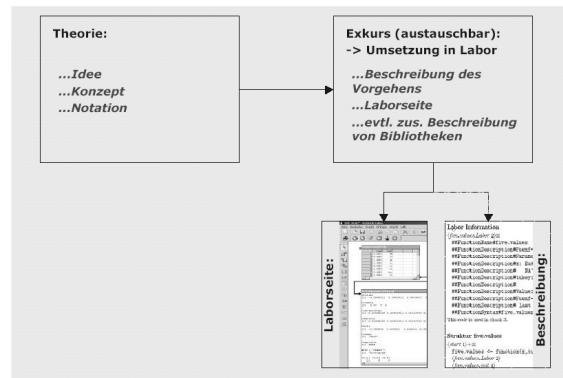


Abbildung 57: Implementation a des Exkurses *Datenanalyse.R*

Der Exkurs sollte in den Theoriekomponenten implementiert werden und von dort abrufbar sein. Er verweist direkt auf Laborseiten und deren Beschreibung. Die Beschreibung enthält eine Erklärung und Dokumentation zu allen in der statistischen Datenanalyse benötigten Objekte, Befehlen und Funktionen. Durch die Implementierung des Exkurses wird die Einbindung und Nutzung von Bibliotheken transparent. Jede Funktion einer Bibliothek wird begründet und ist somit für den Nutzer nachvollziehbar. Er hat die Möglichkeit, Funktionen nach eigenen Bedürfnissen zu modifizieren und neu einzubinden.

Die vorgesehene Implementierung des Exkurses in die Theoriekomponenten wurde durch einen Projektmehrheitsbeschluss verhindert. Dies kann auf die mangelnde Auseinandersetzung mit Umgang, Anwendung und Möglichkeiten des Labors zurückgeführt werden. Der Exkurs wurde folglich in jedes Lernmodul im Bereich *Materialien* implementiert. Damit wurde eine Rolle des Exkurses verhindert. Betrachten wir eine exemplarische Seite des implementierten Exkurses:

<sup>264</sup>Durch die Bezeichnung *Exkurs* wird die Auslagerung der entsprechenden Inhalte impliziert. Diese Auslagerung wurde durch einen Projektbeschluss herbeigeführt. Der Exkurs wird hier allerdings als eine der wichtigsten Komponenten angesehen.

Abbildung 58: Implementation b des Exkurses *Datenanalyse.R*

Die Funktion *KontHaeuf*, ihr Ziel und Aufruf im Statistiklabor werden erklärt. Es findet eine Verknüpfung mit einer Demonstrationsseite im Statistiklabor statt. Auf dieser wird die Anwendung der Funktion *KontHaeuf* an einem Beispiel demonstriert. Zudem werden Hinweise zum Umgang mit der Funktion gegeben. Auf der folgenden Abbildung wird ein weiterer Aspekt des Exkurses deutlich: der Aufruf des Steckbriefes:

Abbildung 59: Steckbrief zu den Funktionen

Zu jeder R-Funktion wird im Exkurs ein Steckbrief bereitgestellt, der deren Aufgaben sowie Umgang und Einsatzmöglichkeiten komprimiert darstellt.

### Erstellung statistischer Reports

Das Erstellen statistischer Reports rückte im Entwicklungsprozess verstärkt in den Mittelpunkt. In der Anfangsphase des Entwicklungsprozesses wurde der Editor implementiert. Der Report entstand durch die Kombination des Editors mit dem Statistiklabor. Diese Lösung war instabil und kompliziert. Objekte mussten kopiert und eingefügt werden. Der Einsatz zeigte, dass eine neue Lösung benötigt wurde. Als Konsequenz wurde die Berichterstellungsfunktion entwickelt und im Labor implementiert. Laborobjekte können für den Report/Bericht ausgewählt werden. Sie werden automatisch in einem Textdokument dargestellt und können in diesem direkt dokumentiert werden. Der Einsatz der Berichterstellungsfunktion hat einen Nachteil aufgedeckt. Bei der Reporterstellung können nur Objekte *einer* Laborseite ausgewählt werden. Wenn Datenanalysen auf mehreren Laborseiten durchgeführt werden, können die ausgewählten Laborobjekte nicht in *einem* Report dargestellt werden.

Arbeitet man als Konsequenz auf nur einer Laborseite, wird die Arbeit schnell unübersichtlich. Dies ist insbesondere auf die Verbindung der Objekte durch Konnektoren zurückzuführen. In Teil I wurden verschiedene statistische Probleme gelöst. Betrachten wir exemplarisch das Beispiel *Strukturen in Daten*. Der Kontext wird in dem folgenden Absatz dargestellt:

„Die Struktur der Daten weist auf einen nicht normalverteilten Charakter hin. Das Ziel besteht darin, den Daten durch eine geeignete Transformation eine annähernd normalverteilte Gestalt zu geben. Zunächst sollen die Studierenden lernen, Strukturen in Daten zu erkennen. Dann sollen sie verstehen, dass die Struktur durch die Transformation der Daten umgewandelt werden kann. Außerdem soll ihnen vermittelt werden, dass eine Transformation von Daten neue Möglichkeiten der statistischen Datenanalyse mit sich bringen kann. Im Anschluss wird die Standardisierung beim statistischen Testen als Transformationsprozess beschrieben.“<sup>265</sup>

Der in dem Beispiel demonstrierte Lösungsprozess findet auf vielen verschiedenen Laborseiten statt.<sup>266</sup>

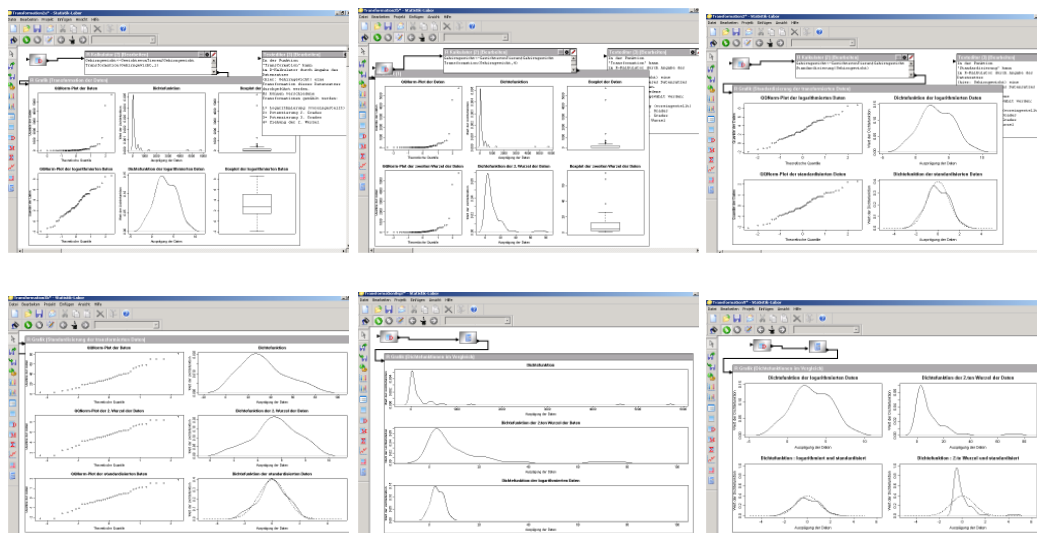


Abbildung 60: Transformationen im Statistiklabor

Es wird deutlich, dass man schnell die Übersicht verlieren würde, wenn alle diese Analysen auf ein und derselben Laborseite durchgeführt würden.

<sup>265</sup> Vgl. 4.3.

<sup>266</sup> In Teil I wurden diese im Detail vorgestellt und interpretiert. Die Details haben hier keine Bedeutung, es soll der Umfang demonstriert werden.

Im Einsatz wurde diese Schwäche der Berichterstellungsfunktion identifiziert, als die Studierenden Reports erstellten. Die Problemstellungen waren umfangreich genug, um während der Datenanalyse zahlreiche Laborobjekte zu benötigen. Die erstellten Laborseiten wurden unübersichtlich. Dies verhinderte die Nachvollziehbarkeit des Lösungsprozesses auf der Laborseite.

In der Weiterentwicklung des Statistiklabors wird das Ziel verfolgt, beliebig viele Laborseiten in einen Report einbinden zu können. In der derzeitigen Endversion des Statistiklabors konnte dieses Ziel technisch noch nicht realisiert werden.

Die Ausführungen haben wiederholt verdeutlicht, wie sich Einsatz- und Entwicklung gegenseitig beeinflussen.

**Der Texteditor als Laborobjekt:** Nach der Entwicklung und Implementierung der Berichterstellungsfunktion erschien der Texteditor zunächst überflüssig. Er wurde folglich aus der Lernsoftware eliminiert. Im Einsatz stellte sich heraus, dass der Texteditor nicht durch die Berichterstellungsfunktion ersetzt werden kann. Der Texteditor ermöglicht eine direkte Dokumentation auf den Laborseiten. Er hat neben der Dokumentation von Problemstellungen und Interpretationen die Aufgabe, den Umgang mit Laborobjekten zu erklären. Betrachten wir zur Verdeutlichung eine Laborseite aus Teil I. Im Texteditor wird der Umgang mit der bereitgestellten Funktion erklärt:

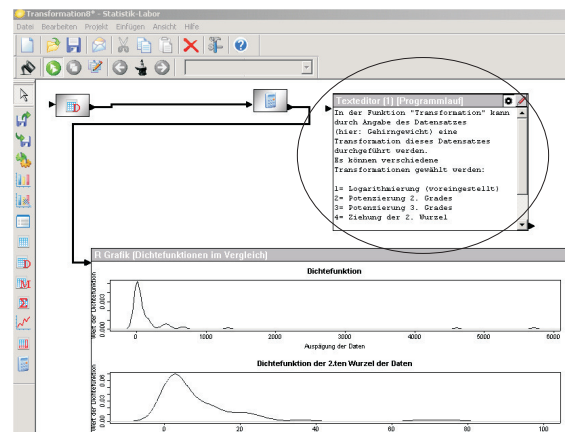


Abbildung 61: Texteditor als Laborobjekt

**Funktionalitäten des Grafik-Wizards:** Der Einsatz zeigte, dass die Studierenden besonders zu Beginn der Labornutzung ein Tool benötigen, das keine R-Kenntnisse erfordert. Der Grafik-Wizard wurde entwickelt, anschließend eingesetzt und positiv von den Studierenden angenommen. Im Einsatz wurden Schwachstellen identifiziert:

- Es war nicht möglich, vergleichende Boxplots zu erzeugen.
- Beim Histogramm konnten keine unterschiedlichen Klassenbreiten und kein Startpunkt gewählt werden.
- Die empirische Verteilungsfunktion war nicht immer im Definitionsbereich, außerdem nahm sie für diskret und stetig skalierte Merkmale die gleiche Form an.

Diese im Einsatz identifizierten Schwachstellen wurden in der sich anschließenden Entwicklungsphase behoben.

Während des Einsatzes wurde der R-Kalkulator eingesetzt, um die Probleme kurzfristig zu beheben. Die Bibliothek *Datenanalyse.R* wurde eingebunden. Sie enthält Funktionen zum Erzeugen vergleichender Boxplots. Außerdem enthält sie eine Funktion, die die Wahl eines Startpunktes und unterschiedlicher Klassenbreiten beim Histogramm ermöglicht. Zudem ist eine Funktion implementiert, die eine Differenzierung zwischen diskret und stetig skalierten Merkmalen bei dem Erzeugen der empirischen Verteilungsfunktion ermöglicht.

**Das Laborobjekt Urliste:** Das Laborobjekt Urliste dient zur manuellen Eingabe eines Datensatzes. Alle hier eingetragenen Werte werden als Vektor repräsentiert. Im Einsatz stellte sich heraus, dass das Objekt Urliste folgende Schwäche hat: Wird es in Verbindung mit dem R-Kalkulator verwendet, ist keine Veränderung des Dateinamens im Nachhinein möglich. Eine Umbenennung führt zum automatischen Löschen aller Einträge der Urliste. Dieser Fehler wurde in der sich anschließenden Entwicklungsphase behoben.

**Das Laborobjekt Kontingenztabelle:** Im Weiterentwicklungsprozess des Statistiklabors wurde das Laborobjekt Kontingenztabelle implementiert. Im Einsatz zeigte sich allerdings, dass dieses nicht genutzt werden kann:

- Das Objekt besitzt neben einem Dateneingang keinen Datenausgang, d.h. dass mit den Daten der Kontingenztabelle nicht weitergearbeitet werden kann.
- Daten können nur aus einer Datei eingelesen, aber nicht direkt in die Kontingenztabelle eingegeben werden. Man kann somit nur mit Daten arbeiten, die in Dateiform vorliegen und alle Beobachtungen enthalten (keine missing-values).
- Bei der Benennung von Spalten/Zeilen der Kontingenztabelle muss jede einzelne Beobachtung benannt werden. Bei kodierten Daten muss z.B. jede "0" zu "männlich" umgewandelt werden, um eine Benennung der dazugehörigen Spalte zu erzeugen.

Der Einsatz führte aus diesen Gründen zu der Erkenntnis, das Objekt Kontingenztabelle vollständig zu überarbeiten. Das Beispiel in Teil I zum Simpson-Paradox demonstriert, wie auch ohne ein lauffähiges Objekt mit Hilfe des R-Kalkulators Kontingenztabellen erzeugt werden können. Diese besitzen zudem den Vorteil, interaktiv veränderbar zu sein.<sup>267</sup>

Die phasenweise Entwicklung von *Statistik interaktiv komplett* wurde anhand der Komponente *Statistiklabor* demonstriert. Die Darstellung ausgewählter Detailentwicklungen des Statistiklabors verdeutlichte, dass Lernmodule und ihr zugehöriges Einsatzkonzept parallel zueinander entwickelt werden müssen. Diese Erkenntnis ist in das Vorgehensmodell in Phase 3 eingeflossen:

- 3a: Entwicklung des Einsatzkonzeptes
- 3b: Entwicklung der Prototypen

Die Ergebnisse der Phase 3 des gewählten Vorgehensmodells wird im Folgenden aus der lerntheoretischen Perspektive betrachtet.

---

<sup>267</sup> Vgl. 4.2.

## 9 Erste Überprüfung der Lernmodule und des Einsatzkonzeptes

”Constructivism - from Philosophy to Practice” ist ein Artikel, der 1997 von Murphy und Rheaume, Universität Laval, Canada im Internet veröffentlicht wurde. Es wird die Verbindung zwischen Konstruktivismus und neuen Möglichkeiten durch neue Technologien hergestellt. Murphy und Rheaume diskutieren, wie konstruktivistische Prinzipien in die Lehre integriert werden können und entwickeln auf Basis dieser Überlegungen eine konstruktivistische Checkliste. Die Checkliste enthält folgende zu überprüfende Charakteristika:

”Multiple perspectives, Student-directed goals, Teachers as coaches, Metacognition, Learner control, Authentic activities and context, Knowledge construction, Knowledge collaboration, Previous knowledge constructions, Problem solving, Consideration of errors, Exploration, Apprenticeship learning, Conceptual interrelatedness, Alternative viewpoints, Scaffolding, Authentic assessment, Primary sources of data.”<sup>268</sup>

Die Checkliste befindet sich auf einem höheren Abstraktionsniveau als die üblicherweise vorgeschlagenen Checklisten zur Beurteilung multimedialer Lernumgebungen. Es werden übergeordnete Aspekte aus der konstruktivistischen Perspektive betrachtet. Übliche Checklisten konzentrieren sich häufig auf technische Details, sind auf einer niedrigeren Betrachtungsebene anzusiedeln. Durch die weit gewählte Begrifflichkeit der konstruktivistischen Checkliste entstehen ein Bewertungs- sowie Gestaltungsspielraum in Bezug auf multimediale Lernumgebungen.

Es wird nun diskutiert, wie die einzelnen Charakteristika in Bezug auf *Statistik interaktiv komplett* und das Einsatzkonzept interpretiert werden können und wie bzw. inwiefern diese in der entwickelten multimedialen Lernumgebung umgesetzt werden konnten. Pro Aspekt wird folgende Differenzierung vorgenommen:

- a) Was bedeutet dieser Aspekt in Bezug auf die Statistikausbildung?
- b) Inwieweit werden die entwickelten Lernmodule und das Einsatzkonzept diesem Aspekt gerecht?
- c) Wie könnten Verbesserungen gestaltet sein? Wie könnte der Aspekt in ein konstruktivistischen Einsatzkonzept oder in die Lernmodule integriert werden?

Es werden drei Aspekte der konstruktivistischen Checkliste diskutiert, die für die Bewertung von *Statistik interaktiv komplett* entscheidend sind.

- Multiple perspectives
- Authentic activities and context
- Problem solving

### Multiple perspectives

a) Statistische Konzepte und Inhalte sollten dem Studierenden aus verschiedenen Perspektiven vorgestellt und vermittelt werden. Durch diese Gestaltung sollen dem Studierenden Zusammenhänge statistischer Methoden verdeutlicht werden. Des Weiteren sollte die Lehre transportieren, dass es oft nicht nur einen Lösungsweg und eine richtige Betrachtungsweise, sondern viele mögliche Wege gibt. Diese Erkenntnis sollte ihn dazu ermutigen, selbst beim Lösen statistischer Probleme verschiedene Sichtweisen auf das Problem zu wählen und sich nicht vorschnell auf einen bestimmten Lösungsweg festzulegen oder nur nach dem einen richtigen zu suchen.

b+c)<sup>269</sup> Die entwickelten multimedialen Lernmodule werden diesem Aspekt z.B. durch ihre Gestaltung gerecht. Das Ziel, Informationen aus Daten zu gewinnen wird aus verschiedenen Perspektiven

<sup>268</sup>[Murphy1997].

<sup>269</sup>Eine Differenzierung bietet sich hier aufgrund der Verzahnung der Inhalte nicht an.



mit Hilfe unterschiedlicher statistischer Methoden verfolgt. Während der Vermittlung der deskriptiven Statistik wird die statistische Datenanalyse aus vier Perspektiven vorgestellt und vermittelt. Im ersten Lernmodul wird ein Datensatz zugrunde gelegt, der nicht durch statistische Methoden komprimiert wird. Es werden statistische Methoden auf die Urliste angewendet und so erste Informationen über die Daten gewonnen (5-Zahlen-Zusammenfassung, Boxplot, empirische Verteilungsfunktion, Maßzahlen). Das zweite Lernmodul konzentriert sich auf die statistische Datenanalyse auf Basis einer nicht klassierten Häufigkeitstabelle. Hier wurden die Daten also bereits komprimiert. Die statistischen Methoden zur Datenanalyse werden aus einem anderen Blickwinkel vermittelt. Zum Teil tauchen andere Methoden auf bzw. die Methoden (Häufigkeitstabelle, Stabdiagramm) des ersten Lernmoduls in modifizierter Form (empirische Verteilungsfunktion aus der Häufigkeitstabelle, Maßzahlen). Im dritten Lernmodul werden die Daten noch stärker komprimiert, die statistischen Methoden und deren Veränderung (Histogramm) auf Basis einer klassierten Häufigkeitstabelle vorgestellt. Im vierten Lernmodul wird die Information der Daten auf eine Zahl (Maßzahlen) reduziert. Die Komprimierung ist vollständig. Unabhängig von der Darstellung des Datensatzes werden Maßzahlen und ihre Eigenschaften diskutiert. Diese Art der Gestaltung vermittelt dem Studierenden statistische Methoden aus unterschiedlichen Perspektiven. Dies zeigen die Startseiten der vier Lernmodule:

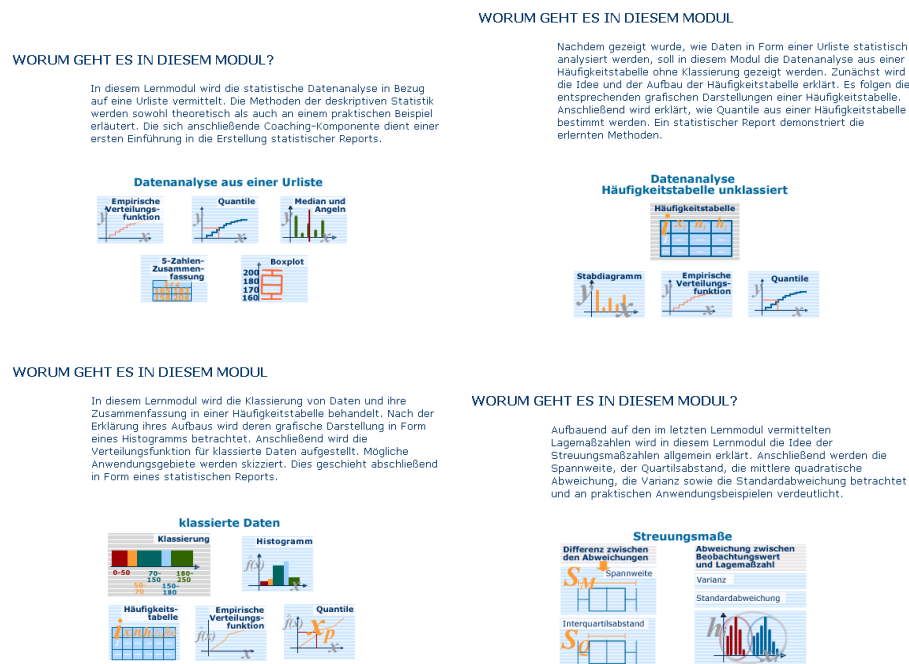


Abbildung 62: Startseiten der Lernmodule - verschiedene Perspektiven

Auch die Vermittlung der statistischen Lehrinhalte innerhalb der Lernmodule erfolgt aus mehreren Perspektiven. Dies wird durch das Vorhandensein unterschiedlicher multimedialer Komponenten innerhalb der Lernmodule erreicht. Erläutern wir dies am Beispiel des Lernmoduls Testen. Die Theoriekomponente vermittelt die theoretischen Aspekte eines statistischen Tests. Die Fallstudie zum Testen erklärt dem Studierenden die Idee statistischer Tests (dies ist in der Theoriekomponente allerdings auch enthalten). Im Statistiklabor wird dem Studierenden die praktische Anwendung des Testens an sich vermittelt. Eine Aufgabe und Musterlösung sind in der Lernumgebung implementiert. Eine Animation erklärt das Testen mit Hilfe von Bild und Ton. Ein Applet ermöglicht dem Studierenden einen Zugang zum Testen durch aktive Nutzung des Applets. Die Übungskomponente stellt ihm ein Problem bereit, anhand dessen er im Statistiklabor durch eigene Aktivität sein erworbenes Wissen festigen und überprüfen kann.

**Authentic activities and context**

a) Statistik sollte anhand realer Daten und Problemstellungen vermittelt werden. Reale Problemstellungen zeigen, wie komplex bestimmte zu lösende Sachverhalte sind. Die Vermittlung statistischer Methoden anhand realer Datensätze zeigt dem Studierenden den direkten Zusammenhang zwischen der Statistik und der Realität.

b) In den entwickelten Lernmodulen wird an vielen Stellen die Brücke zwischen der Realität als dem praktischen Anwendungsgebiet der Statistik und den in der Ausbildung zu vermittelnden statistischen Methoden geschlagen. Statistische Methoden werden anhand realer Datensätze vermittelt. Dies gilt für die Beispiele, Übungen oder Reports. Die Gestaltung der multimedialen Lernumgebung ermöglicht die nachträgliche Ergänzung realer Datensätze durch den Dozenten. Die Modifizierbarkeit soll den langfristigen zukünftigen Einsatz der Lernumgebung gewährleisten.

c) Ergänzt werden diese Aspekte durch die Integration aktueller Datensätze im entwickelten Einsatzkonzept, z.B. werden in der Vorlesung und Übung aktuelle Datensätze der Europäischen Union im Statistiklabor analysiert und in Reportaufgabenstellungen für die Studierenden integriert. Sie wenden somit innerhalb der Prüfung statistische Methoden auf aktuelle Datensätze an.

**Problem solving**

a) Den Studierenden sollen nicht die statistischen Methoden allein, sondern zudem statistische Problemlösungskompetenz vermittelt werden. Unter statistischer Problemlösungskompetenz verstehen wir die Fähigkeit, ein sich stellendes reales Problem statistisch zu formulieren, durch Anwendung geeigneter statistischer Methoden schrittweise zu lösen, diese Lösungsschritte zu begründen und zu dokumentieren und die statistische Lösung auf die Realität zu übertragen. Der Studierende soll durch dieses Vorgehen lernen, dass es notwendig ist, von der Realität in die Modellebene zu wechseln, dort eine Lösung mit Hilfe statistischer Methoden zu erarbeiten und zurück in die Ebene der Realität zu wechseln. Die Dokumentation seiner Lösungsschritte soll auch Alternativen offenlegen, die er während des Problemlösungsprozesses verworfen hat, und diese begründet darstellen. Auf diese Weise lernt der Studierende, transparent zu arbeiten und seine Lösungsschritte für Dritte nachvollziehbar zu dokumentieren.

b) In den Lernmodulen sind mehrere Komponenten integriert, die diesen konstruktivistischen Aspekt umsetzen. Dabei wird zwischen drei Ebenen unterschieden. Bei der Aufgaben- und der Reportkomponente ist eine Differenzierung vorzunehmen. Sie können ohne und mit Musterlösung in der Lernsoftware bereitgestellt werden. Die Entscheidung liegt beim Dozenten, der dies im Lernmanagementsystem einstellen kann. Durch diese Unterscheidung können die Komponenten *Aufgaben* und *Report* im Einsatz auf drei Ebenen in Erscheinung treten:

- Ebene 1: Die integrierten Komponenten Beispiel, Report (mit Musterlösung), Generalbeispiel demonstrieren dem Studierenden die eben beschriebene Vorgehensweise des statistischen Problemlösens. Diese Komponenten erfordern keine Aktivität des Studierenden, sondern präsentieren statistisches Problemlösen.
- Ebene 2: Die Komponenten Aufgaben (ohne Musterlösung) und Übungen hingegen erfordern Aktivität seitens des Studierenden. Statistische Probleme, eingebettet in einen realen Kontext, werden zur Lösung bereitgestellt. Der Studierende wird durch eine Verlinkung darauf aufmerksam gemacht, die Lösung im Statistiklabor anzufertigen. Die Aufgabenstellung beider Komponenten gibt die anzuwendenden statistischen Methoden sowie deren Reihenfolge zur Lösung des Problems bereits vor.
- Ebene 3: Dies ist in der Komponente Report (ohne Musterlösung) nicht der Fall. Ein reales Problem wird formuliert. Die Entscheidung, welche statistischen Methoden zur Lösung angewendet werden sollten, liegt beim Studierenden. Über Methodenauswahl, Reihenfolge der Anwendung und Begründung des eingeschlagenen Lösungsweges entscheidet der Studierende allein. Er kann zeigen, ob er die hinter den statistischen Methoden stehenden Ideen und

Konzepte verinnerlicht hat, ob er in der Lage ist, die Methoden geeignet anzuwenden und außerdem bestehende Zusammenhänge zwischen den einzelnen Methoden herzustellen.

c) In den Einsatzkonzepten wird gezeigt, dass während des Semesters als Prüfungsleistungen weitere Reports gestellt werden können. So kann schrittweise der Lernfortschritt des Studierenden, d.h. sein Lernerfolg bezüglich statistischen Problemlösens identifiziert werden. Die Fehler der Studierenden lassen den Dozenten erkennen, wo Verständnisschwierigkeiten liegen. Auf diese kann er dann direkt in der Lehre reagieren.

**Zusammenfassung:** Wie auch Murphy kommen wir zu dem Schluss, dass es möglich ist, konstruktivistische Elemente effektiv und wirkungsvoll in die neuen Medien zu integrieren. So kann den Defiziten in der Statistikausbildung positiv entgegengewirkt werden. Murphy zieht das folgende Fazit:

”What is clear, however, is that the projects do indeed provide a direct link between theories of knowledge, on one hand, and teaching and learning practices on the other. In each case, constructivist characteristics were clearly evident and, in fact, provided the back bone for and dictated the types of activities, roles and relationships, resources, and types of thinking involved in each project or environment.”<sup>270</sup>

Der Konstruktivismus gibt keine Regeln und Ratschläge zur praktischen Umsetzung und ist weit auslegbar. Diese Eigenschaften spiegeln sich in der Checkliste wider. Die Tatsache, dass wir bei der Anwendung der konstruktivistischen Checkliste auf *Statistik interaktiv komplett* Zuordnungs- und Differenzierungsprobleme hatten, sollte nicht zur Konsequenz haben, eine Modifikation der Checkliste vorzunehmen. Die Modifikation in dem Sinne der Zusammenfassung einiger konstruktivistischer Charakteristika könnte zur Folge haben, dass Informationen verloren gehen, wenn sie auf andere multimediale Lernumgebungen angewendet werden. Die Schwierigkeit der eindeutigen Zuordnung einzelner praktischer Umsetzungen zu konstruktivistischen Charakteristika der Checkliste scheint für die Bewertung dieses Gegenstandes keine entscheidende Rolle zu spielen.

Ein weiterer entscheidender Aspekt ist durch die Diskussion einiger Aspekte der Checkliste deutlich geworden: die Notwendigkeit und tragende Rolle der Einsatzkonzepte. Die Einsatzkonzepte können, wie gezeigt wurde, konstruktivistische Elemente verstärken oder überhaupt erst realisieren.

## 10 Kritische Reflexion zum Forschungsprojekt *Neue Statistik*

Zu Beginn des Projektes war ein Großteil der Projektpartner der Meinung, die bereits existierende Multimediasoftware *Statistik interaktiv!* prinzipiell übernehmen zu können. Mit geringem Zeit- und Arbeitsaufwand sollte sie in *Statistik interaktiv komplett* integriert werden. Die neuen Lernmodule sollten nach derselben Architektur gestaltet werden. Während der oben dargestellten Phasen des Prototypings stellte sich heraus, dass diese Vorstellung illusionär war. Die Projektarbeit krankte eindeutig an fehlender Diskussion. Durch Diskussionen hätten frühzeitig die Prinzipien, Grundideen, Ziele und Gestaltungsfragen des Projektes festgelegt werden können. Die Ausführungen haben gezeigt, dass die Entwicklung einer Lernsoftware erst beginnen kann, wenn das Konzept steht. Dies hat im Projekt *Neue Statistik* nicht stattgefunden. Ausgewählte Problembereiche werden im Folgenden diskutiert.

**Realitätsbezug** Ein Defizit der traditionellen Statistikausbildung besteht in dem fehlenden Bezug zur Realität. Die neuen Medien ermöglichen den Studierenden und Dozenten, die Statistik anhand realer Datensätze zu behandeln. Deren Umfang stellt für die neuen Medien (im Gegensatz zur traditionellen Lehre) kein Problem dar. Der Dozent hat die Möglichkeit, anhand aktueller Themengebiete die Studierenden zur Anwendung statistischer Methoden zu motivieren. Er kann entsprechende Datensätze problemlos in die multimediale Veranstaltung integrieren. In der Komponente *Statistiklabor* können Daten beliebigen Umfangs eingelesen und analysiert werden. In

<sup>270</sup>[Murphy1997].

der rechnergestützten Lehre war die Verbindung zwischen Realität und Statistik auch möglich<sup>271</sup>, stellte für die Studierenden aber häufig eine Hürde dar. Sie benötigten Programmierkenntnisse, um Datensätze am Rechner zu analysieren. Das *Statistiklabor* (als Komponente der multimedialen Lernsoftware) ermöglicht es den Studierenden, statistische Datenanalysen durchzuführen, ohne Kenntnisse in einer Programmiersprache zu besitzen. Sie bekommen mit dem *Statistiklabor* ein Tool an die Hand, das es ihnen ermöglicht, ein Gefühl für die Daten zu bekommen. Die neuen Medien scheinen in diesem Aspekt sowohl der traditionellen als auch der rechnergestützten Statistikausbildung gegenüber Vorteile zu besitzen.

**Gewichtung von Komponenten** Der Entwicklungsprozess des Statistiklabors wurde bereits diskutiert. Die Implementierung des Exkurses stellte die notwendige Verbindung zwischen den Theoriekomponenten und der Umsetzung der statistischen Konzepte im Statistiklabor her. Auch die Beschreibung von Bibliotheken wurde im Exkurs vorgenommen. Die Implementierung des Exkurses stieß trotz dieser Aufgaben auf erheblichen Widerstand bei vielen Projektpartnern. Es wurde nicht erkannt, dass der Exkurs notwendig für den zielgerichteten Einsatz des Statistiklabors ist. Als die Idee des Exkurses vorgetragen und zur Diskussion gestellt wurde, wurde keine Diskussion ermöglicht. Letztendlich wurde der Exkurs implementiert, allerdings sehr viel weiter im Hintergrund, als dies zunächst vorgesehen war.

**Flexibilität durch Modifizierbarkeit** Bei der Entwicklung der neuen Lernsoftware stand folgende Idee im Vordergrund. Der Dozent soll die Möglichkeit bekommen, seine Veranstaltung individuell gestalten zu können. Er soll eigene Materialien (z.B. Daten) problemlos in die multimedialen Lernmodule integrieren können. Wird die Lernsoftware im Rahmen eines Lernmanagementsystems<sup>272</sup> eingesetzt, besteht die Möglichkeit, eigene Laborseiten zu erstellen und in der Vorlesung zu behandeln. Die bereitgestellten Laborseiten können nach individuellen Vorstellungen modifiziert werden. Angefertigte Dokumentationen zu jeder Laborseite erleichtern dies. Der Dozent kann jederzeit nachvollziehen, was auf der Laborseite geschieht, kann sie verstehen und gezielt verändern oder ergänzen. Dies war allerdings auch in der rechnergestützten Lehre möglich. Kritisch ist an dieser Stelle die Bedeutung des Lernmanagementsystems zu sehen, ohne das die dargestellte Modifizierbarkeit nicht besteht. Die Lernsoftware allein ermöglicht keine Modifikation existenter Lernmodule bzw. Statistiklaborseiten.

Bezüglich der Modifizierbarkeit der multimedialen Lernmodule wird über eine Version nachgedacht, in der der Dozent Theoriekomponenten neu gestalten und ergänzen, Beispiel- oder Aufgabenkomponenten durch eigene austauschen kann. Es soll ihm ein Lager an multimedialen Lehrmaterialien bereitgestellt werden, aus dem er einzelne Komponenten oder auch ganze Lernmodule beliebig auswählen, modifizieren oder durch eigenes Material ergänzen kann.

Es gab Versuche in der rechnergestützten Lehre, Teile dieses Aspektes umzusetzen. Die neuen Medien unterscheiden sich von diesen Versuchen durch das Angebot eines Lernmanagementsystems, das es ermöglicht, viele verschiedene Teile einer Gesamtveranstaltung als stimmiges Ganzes anzubieten. Wieder scheinen die neuen Medien gegenüber der traditionellen und rechnergestützten Statistikausbildung Vorteile zu bieten. Die Lösung dieses Problems über ein Lernmanagementsystem kann nur als Zwischenlösung betrachtet werden, da Lernmanagementsysteme teuer in ihrer Nutzung sind. Dies steht einem flächendeckenden, dauerhaften Einsatz der Lernsoftware im Weg.

**Erhöhtes Maß an Aktivität/Interaktivität** Als ein Defizit der traditionellen Statistikausbildung wurde die geringe Aktivität der Studierenden herausgestellt.<sup>273</sup> Das *Statistiklabor* soll zum aktiven Arbeiten mit Daten motivieren. Die Benutzeroberfläche enthält verschiedene selbsterklärende Buttons, mit deren Hilfe der Studierende durch die Datenanalyse geführt wird. Ein *Grafik-Wizard* ermöglicht ihm die Auswahl gewünschter Diagramme durch Mouseklick. Der Studierende hat die Möglichkeit, auf diese Art und Weise eine komplexe Datenanalyse durchzuführen, ohne zu

---

<sup>271</sup>Analyse realer Datensätze

<sup>272</sup>Die Idee, Aufgabe und Nutzung von Lernmanagementsystemen wird im Rahmen der Evaluationsstudie im sich anschließenden Teil IV dieser Arbeit ausführlich diskutiert.

<sup>273</sup>Vgl. 3.

programmieren.<sup>274</sup> Dies hat sich als eine deutliche Verbesserung gegenüber der rechnergestützten Lehre herausgestellt. Der Studierende erlernt die statistischen Methoden, ihre Ideen und Sensitivitäten durch eigene praktische Anwendung. Er wird zum Lernen von Statistik motiviert, indem er diese praktisch zur Lösung von Problemen anwendet. Ein weiterer Vorteil der Arbeit im Statistiklabor besteht darin, dass der Studierende die Arbeitsschritte des Statistikers aktiv erleben und nachvollziehen kann. Die Laborobjekte sind so entwickelt und miteinander kombinierbar, dass sie die Arbeitsweise des Statistikers widerspiegeln.

Des Weiteren wird die Interaktivität zwischen Studierenden und Dozenten durch den Einsatz der Lernsoftware im Rahmen eines Lernmanagementsystems gefördert. Erstellte Laborseiten können von beiden Seiten bereitgestellt und abgerufen werden. Durch die Möglichkeit, Vorlesungslaborseiten selbst zu starten, kann der Studierende interaktiv im Statistiklabor alle Datenanalysen nachträglich selbst durchführen und so ein Gefühl für Daten und statistische Methoden bekommen. Es besteht die Möglichkeit, den Studierenden durch Hyperlinks von diesen Laborseiten auf Theorie-, Beispiel- oder Aufgabenkomponenten zu verweisen, die die Methoden der entsprechenden Laborseite aufgreifen und geeignete Trainingsmöglichkeiten<sup>275</sup> bieten. Der Studierende hat durch diese Kombination viele Möglichkeiten, die Statistik durch eigene Aktivität zu erlernen.

Auch die Interaktivität zwischen den Studierenden wird erhöht. Die einzelnen Studierenden oder Gruppen können innerhalb des Lernmanagementsystems Lösungen und Aufgaben austauschen. Dies macht sie unabhängig von Zeit und Raum und erleichtert ihnen die Kommunikation untereinander. Sie können z.B. gemeinsam eine Aufgabe im Labor lösen. Die einzelnen Lösungsschritte können innerhalb des Statistiklabors für den anderen bereitgestellt werden. Durch die Berichterstellungsfunktion können sie ihre Analyseergebnisse für den anderen direkt dokumentieren.

Die Lernsoftware *Statistik interaktiv komplett* ermöglicht die dargestellten Interaktivitäten ohne die Nutzung eines Lernmanagementsystems nicht. Bei einem Großteil der Projektpartner bestand allerdings kein Diskussionsbedarf bezüglich dieser Aspekte. Die Notwendigkeit der Entwicklung einer Lernsoftware, die den Aspekten Modifizierbarkeit und Interaktivität genügt, wurde nicht erkannt.

**Reporterstellung:** In der traditionellen Statistikausbildung wurde bemängelt, dass die Studierenden zwar die statistischen Methoden beherrschten, jedoch nicht in der Lage waren, statistisch zu arbeiten, d.h. statistische Reports zu verfassen. Die neuen Medien bieten viele technische Möglichkeiten, Datenanalyse und Dokumentation miteinander zu verbinden. In der entwickelten multimedialen Lernsoftware *Statistik interaktiv komplett!* ist zu diesem Zweck die *Berichterstellungsfunktion* integriert. Diese ermöglicht es dem Studierenden, seine Datenanalyseergebnisse aus dem Statistiklabor zu dokumentieren, zu interpretieren und zu erläutern.

In der traditionellen Lehre konnten größere Datensätze nicht analysiert werden, die Reporterstellung wurde nicht vermittelt. Die neuen Medien bieten die Möglichkeit, dem Studierenden die Erstellung statistischer Reports als Arbeitserleichterung zu vermitteln. Er kann seine Datenanalyseergebnisse schnell und problemlos kommentieren. Er lernt das Verfassen statistischer Reports während der Lösung von Aufgaben oder des Nachvollziehens von Vorlesungslaborseiten oder Beispielen. Ein Problem können die neuen Medien allerdings nicht lösen. Der Arbeitsaufwand des Dozenten steigt durch die Korrektur statistischer Reports enorm. In *Statistik interaktiv komplett* ist die Reportvermittlung stark verankert:

- Die Coachingkomponente erklärt die Idee und die Vorgehensweise der Reporterstellung.
- Musterreports geben dem Studierenden Beispiele zur Reporterstellung.
- Im Statistiklabor stehen Analyse- und Dokumentationsinstrumente bereit. Die Berichterstellungsfunktion bildet den Rahmen der Reporterstellung.

Der Austausch selbst erstellter Reports ist jedoch nur in Kombination mit einem Lernmanagementsystem möglich.

---

<sup>274</sup> Dass dies nicht völlig ausgeklammert werden muss, haben die Betrachtungen des Statistiklabors gezeigt. Es besteht aber die Möglichkeit, Daten zunächst ohne den Einsatz von R-Befehlen zu analysieren.

<sup>275</sup> Trainingsgegenstand ist die Statistik.

**Visualisierungen:** Neue Medien ermöglichen neue Wege der Wissensvermittlung. In die multimediale Lernsoftware werden Animationen und Applets integriert. Diese verdeutlichen einige statistische Methoden und zeigen dem Studierenden Anwendungsmöglichkeiten. Animationen eignen sich zwar lediglich zum passiven Betrachten, Applets hingegen wurden in dem Projekt *Neue Statistik* so gestaltet, dass der Studierende aktiv wird. Für *Statistik interaktiv komplett* kann zusammenfassend festgehalten werden. Die Idee der Entwicklung, Implementierung und Verwendung von Animationen und Applets ist gut, die Umsetzung schlecht. Betrachten wir zwei Beispiele interaktiver Applets, um diese Behauptung zu belegen:

Appletbeispiel 1 - Grafische Darstellung von qualitativen Merkmalen: Der Studierende hat die Möglichkeit, einen von zwei Datensätzen auszuwählen. Eigene Datensätze können nicht eingebunden werden. Modifizierbarkeit oder individuelle Gestaltungsmöglichkeiten werden von dem Applet somit nicht unterstützt.

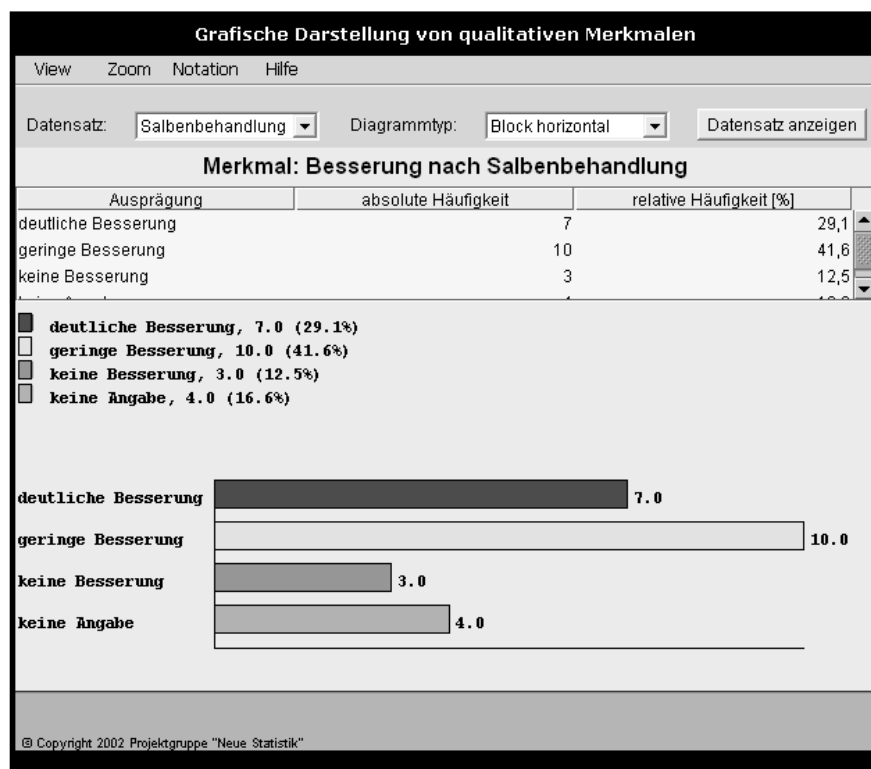


Abbildung 63: Prototyp II: Appletbeispiel 1

In dem Applet erfolgt eine grafische Darstellung der Häufigkeiten eines Datensatzes. Es kann zwischen verschiedenen Diagrammtypen ausgewählt werden: Kreisdiagramm, Stabdiagramm, Säulendiagramm in horizontaler oder vertikaler Sicht. Außerdem besteht die Möglichkeit, sich die entsprechende Häufigkeitstabelle anzeigen zu lassen. Man kann sich auch nur die Häufigkeitstabelle anzeigen, d.h. die Grafik ausblenden lassen. Des Weiteren kann eine detaillierte Betrachtung der beiden zur Verfügung stehenden Datensätze erfolgen. Die Erstellung der Grafik erfolgt in verschiedenen Farben, eine Legende erklärt diese. Die Erklärung zur verwendeten Notation steht im Hintergrund bereit. Durch die integrierte Hilfsfunktion kann sich der Studierende die Ziele des entsprechenden Applets, die Handhabung und die Erklärung einzelner Funktionalitäten anzeigen lassen.

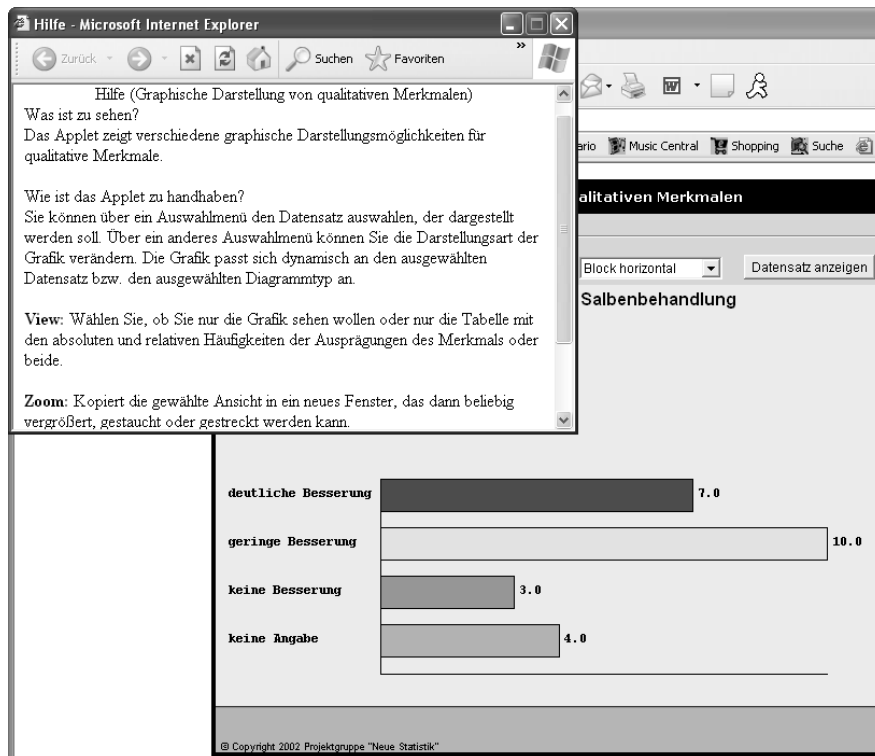


Abbildung 64: Prototyp II: Appletbeispiel 2

Appletbeispiel 2 - Raten von Korrelationen: In diesem Applet erscheinen vier verschiedene Streudiagramme (Plot A - D), die unterschiedliche Punktwolken beinhalten. Unterhalb dieser Streudiagramme sind vier Korrelationskoeffizienten angegeben.<sup>276</sup>

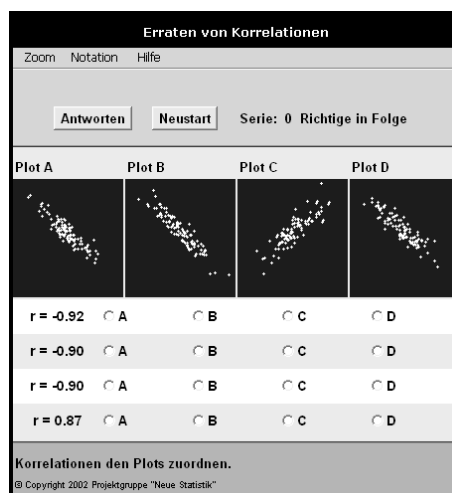


Abbildung 65: Prototyp II: Appletbeispiel 3

<sup>276</sup>Dieses Applet bedarf der Verbesserung: Die Wahl der Korrelationskoeffizienten ist ungeeignet.

Die Interaktionsmöglichkeit des Studierenden besteht darin auszuwählen, welcher Korrelationskoeffizient zu welchem Plot gehört. Nach der Zuordnung klickt er auf ein Feld **Antworten** und bekommt angezeigt, inwiefern seine Antworten zutreffen. Durch die Betätigung der Buttons **Neustart** und **neue Plots** erscheinen vier neue Plots mit neuen Korrelationskoeffizienten. Dieser Vorgang kann vom Studierenden zu Übungszwecken beliebig häufig wiederholt werden. Auch dieses Applet bietet die Möglichkeit, sich Erklärungen zur verwendeten Notation anzeigen zu lassen. Die Applets können durch die Funktion **Zoom** in einem neuen Fenster neu erstellt werden. In diesem ist es größenveränderlich und somit angenehmer zu bearbeiten. Die verankerten Ideen können positiv bewertet werden. Doch auch in diesem Applet ist die Umsetzung schlecht, was durch die Abbildung des Applets deutlich wird: Der Studierende soll unterscheiden zwischen  $r=-0.92$ ,  $-0.90$ ,  $-0.90$  und  $0.87$ . Die Korrelationen hätten sich stärker unterscheiden müssen, um Unterschiede in den Plots identifizieren zu können.

**Zusammenfassung des Teils III** Zunächst wurde das Prototyping als Methode der Softwareentwicklung beschrieben. Das gewählte Vorgehensmodell, insbesondere die Notwendigkeit eines iterativen Entwicklungsprozesses wurde diskutiert. Die Schwerpunkte und Ziele der drei Entwicklungsphasen von *Statistik interaktiv komplett* wurden dargestellt. Am Beispiel der Komponente *Statistiklabor* wurden die phasenweise Entwicklung und die Parallelität von Einsatz und Entwicklung demonstriert. Anschließend wurde auf Basis einer konstruktivistischen Checkliste überprüft, inwiefern die entwickelten Lernmodule und das Einsatzkonzept den Prinzipien des Konstruktivismus gerecht werden. In diesem Rahmen wurden verschiedene Komponenten der Lernmodule vorgestellt. Durch eine kritische Reflexion über Inhalte und Prozesse des Projektes *Neue Statistik* wurde dem Leser abschließend ein Blick hinter die Fassaden des Entwicklungsprozesses ermöglicht.

Im folgenden Kapitel wird in Form einer Evaluationsstudie empirisch überprüft, ob und inwiefern die entwickelten Komponenten der Lernsoftware und des Einsatzkonzeptes dazu führen können, den Lernerfolg der Studierenden positiv zu beeinflussen und Defizite in der Statistikausbildung entgegenzuwirken. Die Ergebnisse der empirischen Überprüfungen werden diskutiert. Es werden Konsequenzen für die zukünftige Entwicklung und den Einsatz statistischer Lernsoftware abgeleitet.



## Teil IV

# Explorative Evaluationsstudie - eine empirische Untersuchung

## 11 Einsatzszenarien, -konzepte und Evaluationsformen

**Überblick** Im Rahmen des Projektes *Neue Statistik* wurde eine Evaluationsstudie konzipiert und durchgeführt, über die im Folgenden reflektierend berichtet wird. Diese bestand aus zwei Durchführungen. Das übergeordnete Ziel der Studie war die schrittweise Evaluation der Lernsoftware *Statistik interaktiv komplett* und des für sie entwickelten Einsatzkonzeptes. Bei beiden Durchführungen wurde explorativ evaluiert. Das explorative Evaluieren setzte sich aus einem aktiven und einem passiven Teil zusammen. Die Studierenden wurden begleitend zum Einsatz der Lernsoftware beobachtet. Diese Methode wird als *teilnehmendes Beobachten* bezeichnet. Der Beobachter blieb passiv. Anschließend fand eine Interaktion mit dem System statt. Der Beobachter griff aktiv in den Vermittlungs- oder Lernprozess ein. Seine Interaktionen fanden als Reaktion auf die Beobachtungen statt. Er steuerte, korrigierte, verbesserte den Lern- und Entwicklungsprozess der Studierenden. Jede Interaktion hatte das Ziel, das System positiv zu verändern. In der ersten Durchführung der Evaluationsstudie wurde ausschließlich qualitativ, in der zweiten vorwiegend quantitativ evaluiert.<sup>277</sup> Die zweite Evaluationsstudie integrierte qualitative Elemente dort, wo die quantitative Evaluation an Grenzen stieß.

**Vorgehensweise** Der Teil IV ist wie folgt aufgebaut. Zunächst erfolgen allgemeine Ausführungen zu Einsatzkonzepten. Es wird gezeigt, dass diese unter Berücksichtigung verschiedener Einsatzszenarien entwickelt werden müssen. Konstruktivistische Einsatzkonzepte werden charakterisiert und ihre Notwendigkeit begründet. Außerdem erfolgt die Abgrenzung zwischen qualitativer und quantitativer Evaluation. Gemeinsamkeiten der Einsatzkonzepte Ia und Ib, die den beiden Durchführungen der Evaluationsstudie zugrunde liegen, werden herausgestellt. Es handelt sich um Varianten des schrittweise entwickelten Einsatzkonzeptes.

Anschließend werden ausgewählte Aspekte der ersten und zweiten Durchführung der Evaluationsstudie vorgestellt und diskutiert. In der ersten Durchführung der Evaluationsstudie wird ein Einsatzkonzept entwickelt und explorativ mit Hilfe qualitativer Methoden evaluiert. Es werden erste Erkenntnisse über den Einsatz multimedialer Lernsoftware in der Statistikausbildung gewonnen und Konsequenzen für die zweite Durchführung der Evaluationsstudie abgeleitet. Die beiden Durchführungen werden dem Leser auf verschiedene Arten präsentiert:

- Aus der ersten Durchführung werden ihm Miniaturen des Evaluationsprozesses beschrieben, ausgewählte *stories* erzählt, die ihm ein Gefühl für die Schwierigkeiten der Evaluation von Multimedia geben. Die Beispiele zeigen, an welche Grenzen die qualitative Evaluation stößt.
- Die zweite Durchführung wird dem Leser strukturiert präsentiert. Er wird in den Entwicklungsprozess eines Evaluationsinstrumentes eingebunden, kann den Prozess schrittweise miterleben.

Während der zweiten Durchführung fand die Entwicklung eines Evaluationsinstrumentes statt. Mit dessen Hilfe können multimediale Lernsoftware sowie ihr Einsatz evaluiert werden. Das Evaluationsinstrument bietet den Einsatz statistischer Analyseverfahren an, konzentriert sich auf die quantitative Evaluation. Die Umsetzung der Analyseverfahren findet auf speziell entwickelten Statistikkaborseiten statt. Die zweite Durchführung der Evaluationsstudie kann als Pretest des entstehenden Evaluationsinstrumentes angesehen werden. Eine quantitative Evaluation wird durchgeführt und in Form eines statistischen Reports dokumentiert. Qualitative Elemente werden in den Report aufgenommen. Diese sind stark abhängig von dem zugrunde liegenden Einsatzkonzept.

---

<sup>277</sup>Die Abgrenzung zwischen qualitativer und quantitativer Evaluation erfolgt in einem der folgenden Unterkapitel.

Stellt die Vermittlung des statistischen Reportverfassens ein Ziel des Einsatzes dar, so müssen die Evaluationsmethoden so gewählt werden, dass die Erreichung dieses Zieles überprüft werden kann. Im Rahmen der zweiten Durchführung wird außerdem gezeigt, unter welchen Umständen die Anwendung von Lernmanagementsystemen vorteilhaft für den Einsatz neuer Medien ist.

Aus den Erfahrungen der Evaluationsstudie wird ein Resümee gezogen. Es erfolgt eine abschließende Bewertung der entwickelten Lernsoftware und des Einsatzkonzeptes.

**Übergeordnete Ziele der Evaluationsstudie:** Das übergeordnete Ziel bestand darin, Erkenntnisse über den Einsatz multimedialer Lernsoftware in der statistischen Grundausbildung zu sammeln. Auf dieser Basis werden Empfehlungen über bestimmte Vermittlungskonzepte, d.h. verschiedene Arten der Integration von Multimedia in Lehrveranstaltungen gegeben. Durch die empirische Überprüfung unterschiedlicher Konzepte wird versucht, Aussagen über Komponenten der Lernsoftware und des Einsatzkonzeptes zu treffen, die den Lernerfolg der Studierenden entscheidend beeinflussen oder die zur Behebung bestimmter Defizite in der Statistikausbildung besser geeignet sind als andere.

Im Rahmen der Studie fand die Beantwortung folgender Fragen statt:

- Wird das Verständnis der Studierenden für statistische Fragestellungen, das Verständnis, reale statistische Probleme zu erkennen und selbstständig zu bearbeiten und zu lösen, durch den Einsatz multimedialer Lernsoftware in der statistischen Grundausbildung gefördert?
- Können bestimmte Einsatz-/Vermittlungsformen den Lernerfolg und die Motivation der Studierenden erhöhen?
- Ist es möglich oder Erfolg versprechend, die statistische Grundausbildung flächendeckend, d.h. in ganz Deutschland mit Studierenden eines ganzen Semesters durch den Einsatz statistischer Lernsoftware zu unterstützen?
- Inwieweit sind die erzielten Ergebnisse auf den Einsatz der Multimediasoftware bzw. bestimmte Einsatz-/Vermittlungskonzepte zurückzuführen?

Die folgenden Unterkapitel werden Antwortversuche geben. Es wird der Versuch unternommen, ein vielseitiges Einsatzkonzept zu entwerfen, das geeignet ist, multimediale Lernsoftware in der Zukunft effizient an deutschen Universitäten einzusetzen.

### 11.1 Charakterisierung: Einsatzkonzepte und Einsatzszenarien

Vorab wird definiert, was unter einem Einsatzkonzept verstanden wird. Wir charakterisieren ein Einsatzkonzept durch zwei Aspekte:

- Materialien
- Umgang mit den Materialien

Die Materialien werden mit dem Ziel eingesetzt, Defiziten der Statistikausbildung entgegenzuwirken. Anschließend wird der Grad der Zielerreichung evaluiert.

Im Rahmen der Evaluationsstudie wurden zwei Varianten eines Einsatzkonzeptes entwickelt und überprüft. Inhaltlich gesehen unterscheidet sich das Einsatzkonzept Ia nicht wesentlich vom Einsatzkonzept Ib, da beide für dasselbe Einsatzszenario, den Einsatz in der statistischen Grundausbildung der Universität, konzipiert sind. Unterschiede und Gemeinsamkeiten werden im folgenden Unterkapitel dargestellt. Ein Unterschied besteht in dem Grad der Betreuung. Bei der Gestaltung des Einsatzkonzeptes Ia stand der Gedanke des *self-paced-learning* im Mittelpunkt. Die Evaluation des Einsatzkonzeptes Ia zeigte, dass das *self-paced-learning* nicht ohne weiteres funktioniert, die Studierenden ohne geeignetes *guiding* große Schwierigkeiten haben, das ihnen zur Verfügung stehende Material effizient zu nutzen.

### Konstruktivistische Einsatzkonzepte für die neuen Medien

Die Wirkung der neuen Medien oder des Rechners an sich darf nicht überschätzt werden. Die Erfahrung lässt uns die These aufstellen, dass der Erfolg von Lernsoftware in der Statistikausbildung zu einem großen Teil vom Dozenten abhängt.

„Der Konstruktivismus, wenn man versucht, ihn als pädagogische Anweisung zu interpretieren, definiert den guten Pädagogen als einen kreativen, moderierenden und Alternativen aufzeigenden solchen, nicht als einen, der lediglich belehrt und glaubt, mit multiple-choice Fragen etwas über seine Studierenden und ihre Sicht auf die Welt sowie deren Erklärung zu erfahren.“<sup>278</sup>

Einsatzkonzepte übernehmen unter anderem die Aufgabe, den Dozenten über seine Möglichkeiten zu informieren. Sie machen ihm bewusst, dass er mit Hilfe der neuen Medien statistische Konzepte durch neue Möglichkeiten anders vermitteln kann. Ein Dozent, der in dem Rechner einen „Rechner“ sieht, kann auch mit der besten Lernsoftware keine Erfolge in der Statistikausbildung erzielen. Wird ihm das Statistiklabor demonstriert, kann er die Vorteile erkennen, statistische Methoden zu vermitteln. Bei der Entwicklung und Bereitstellung eines konstruktivistisch geprägten Einsatzkonzeptes standen folgende Fragen im Mittelpunkt:

- Was zeichnet eine konstruktivistische Statistikausbildung aus? Welche konstruktivistischen Prinzipien sollen in das Einsatzkonzept aufgenommen werden?
- Welche Möglichkeiten bestehen durch die neuen Medien, konstruktivistische Prinzipien in die Statistikausbildung zu integrieren?
- Wie können traditionelle, multimediale und konstruktivistische Elemente geeignet miteinander kombiniert werden?<sup>279</sup>
- Kommt dem Dozenten durch das Einsatzkonzept eine neue Rolle zu, muss diese genau beschrieben werden. Welche Anweisungen müssen im Einsatzkonzept festgehalten werden, um den Dozenten zu einem konstruktivistisch geprägten Statistikerunterricht zu leiten?

Der Autor vertritt die Meinung, dass ein Einsatzkonzept durch sein Einsatzszenario bestimmt wird.

**Existenz verschiedener Einsatzszenarien** Die entwickelte Lernsoftware *Statistik interaktiv komplett* könnte in den folgenden Szenarien eingesetzt werden:

1. Einsatz in der Universität mit den Veranstaltungskomponenten Vorlesung, Übung, Rechnerstunden. Es wird gezeigt, dass in diesem Szenario verschiedene Variationen des Einsatzes stattfinden können. Je nach Zielsetzung des Dozenten kann wie bereits erwähnt eines der konstruktivistischen Prinzipien, *self-paced-* oder *guided learning*, stärker verfolgt werden.
2. Fernstudium
3. Eigenständig Lernender: unabhängig von der Universität, ohne externe Betreuung

Jedes Einsatzszenario ist mit bestimmten Bedingungen verbunden, setzt Gegebenheiten voraus, durch die es sich von den anderen Szenarien unterscheidet. In der durchgeführten Evaluationsstudie wurde das erste Einsatzszenario exemplarisch überprüft. Diese Wahl ist durch die Ausrichtung des Forschungsprojektes *Neue Statistik* begründet. Es wurde das Ziel verfolgt, statistische Lernsoftware zu entwickeln, die flächendeckend an deutschen Hochschulen eingesetzt werden kann.<sup>280</sup> Dass innerhalb des Einsatzszenarios verschiedene Gestaltungsmöglichkeiten bestehen, zeigen die Varianten Ia und Ib des entwickelten und evaluierten Einsatzkonzeptes.

<sup>278</sup>[Tiemann2002], S.180

<sup>279</sup>Mögliche Kombinationsmöglichkeiten werden dem Leser demonstriert. In dem Zusammenhang werden Unterschiede und Gemeinsamkeiten deutlich.

<sup>280</sup>Dieses Ziel geht konform mit dem der vorliegenden Arbeit, Defizite in der Statistikausbildung zu beheben.

Unabhängig von dieser Vorgabe dominiert das erste Einsatzszenario in der Realität beide anderen. Es gibt deutlich weniger Fernstudierende und eigenständig Lernende als Studierende deutscher Hochschulen. Auch aus diesem Grunde wurde das erste Einsatzszenario ins Zentrum der Evaluationsstudie gestellt. Um die Fülle der von verschiedenen Dozenten vertretenden Interessen zu berücksichtigen, müssen Lernmodule entwickelt werden, die folgenden Operationen ermöglichen:

- Wiederverwendbarkeit
- Austausch
- Selektion

Für diese Operationen müssen die Lernmodule feste und variable Teile enthalten. Dies wirkt sich unter anderem auf die Gestaltung der Texte aus. Sind sie zu eng aufeinander abgestimmt, können einzelne Komponenten bzw. Module nicht problemlos herausgegriffen oder neue integriert werden. Es stellt sich die Frage, wie elastisch das Netz von Lernmodulen und Komponenten sein muss, um ein fast beliebiges Herausgreifen einzelner Komponenten zu ermöglichen.

In den einzelnen Durchführungen der Evaluationsstudie werden Vorschläge unterbreitet, wie man zur Erreichung bestimmter Ziele mit den Lernmodulen umgehen kann. Dies erfolgt exemplarisch für das ausgewählte Einsatzszenario. Es wird deutlich, dass die entwickelten Lernmodule Schwächen aufweisen, d.h. bestimmte Operationen nicht ermöglichen. Dies kann auf die Ausgangssituation des Projektes *Neue Statistik* zurückgeführt werden. Die verantwortlichen Professoren hatten Erfahrung in der Statistikausbildung und besaßen somit eine klar definierte individuelle Bedürfnisstruktur. Diese war nicht immer deckungsgleich mit der Bedürfnisstruktur, die in den Durchführungen der Evaluationsstudie identifiziert wurde. Mangelnde Kommunikations- und Kompromissbereitschaft führten dazu, dass der identifizierte Bedarf, die Anforderungen an Lernmodule und Einsatzkonzept nicht vollständig im Entwicklungsprozess umgesetzt wurden. Dies führte zu Schwachstellen der entwickelten Materialien.

Der Idealzustand könnte wie folgt aussehen. Der Einsatz und die Evaluation der Prototypen identifiziert nicht befriedigte Bedürfnisse der Dozenten und Studierenden. In der sich anschließenden Phase des Prototypings werden diese aufgegriffen und bestmöglich behoben. Dies geschieht durch eine Modifikation der Lernmodule und Komponenten des Einsatzkonzeptes.

Bevor die beiden Durchführungen der Evaluationsstudie beschrieben und diskutiert werden, wird die hinter der Entwicklung konstruktivistischer Einsatzkonzepte stehende Intention erklärt. Außerdem wird der durch die Ziele in der Statistikausbildung aufgespannte Raum näher beschrieben.

## 11.2 Assessment in Statistical Education

Die Neugestaltung der Statistikausbildung ist geprägt durch Veränderungen in der Vermittlung und Bewertung der Leistungen der Studierenden. Garfield spricht von "assessment revolution"<sup>281</sup>. Betrachten wir seine Ausführungen im Folgenden genauer, um Parallelen zu den entwickelten Lernmodulen und dem Einsatzkonzept zu ziehen. Viele Ideen wurden aufgegriffen und in die Gestaltung des Einsatzkonzeptes integriert.

*Assessment* hat seiner Meinung nach eine neue Bedeutung bekommen:

"Newer definitions of assessment refer to procedures more likely to elicit complex intellectual performance than traditional paper and pencil tests. Some of these alternative procedures now include portfolios of student's work, students projects, and laboratory activities"<sup>282</sup>

Statt wie in der traditionellen Lehre Fähigkeiten losgelöst vom Problem zu überprüfen, konzentrieren sich die neuen Methoden darauf, zu überprüfen, ob die Studierenden statistische Konzepte verstanden haben. Sie testen, ob Studierende in der Lage sind, statistisches Wissen auf die Lösung

---

<sup>281</sup>[Garfield1999].

<sup>282</sup>Darling-Hammond et al., 1995 in [Garfield1999].

von Problemen anzuwenden und außerdem in der Sprache der Statistik zu kommunizieren. Der Wandel in der Statistikausbildung wird durch zahlreiche neue Bewertungs- und Vermittlungsmethoden vollzogen, die sich auf die Gestaltung des Statistikerunterrichts auswirken.

Garfield and Gal formulierten Prinzipien des Statistikerlernens. Diese stellen die Ziele dar, die wir in der Statistikausbildung erreichen wollen:<sup>283</sup>

- "Understand the purpose and logic of statistical investigations
- Understand the process of statistical investigations
- Master important procedural skills
- Understand probability and chance
- Develop interpretative skills and statistical literacy
- Develop ability to communicate statistically
- Develop useful statistical dispositions"

Betrachten wir zusätzlich die von Garfield als wichtig erachteten Grundpfeiler konstruktivistischer Didaktik. Es wird für jeden exemplarisch dargestellt, wie sie sich im Einsatzkonzept niederschlagen:<sup>284</sup>

- *Students learn by constructing knowledge*: Die statistischen Reports verlangen die Kombination verschiedener statistischer Methoden. Die Studierenden verknüpfen altes und neues Wissen miteinander, um statistische Probleme zu lösen, konstruieren durch diese Prozesse neues Wissen.
- *Students learn by active involvement in learning statistics*: Die statistischen Reports werden im Statistiklabor erstellt. Die Studierenden werden im Labor aktiv, führen dort die Analyse von Daten durch. Auch in der Übung präsentieren sie ausgewählte Lösungen am Rechner, arbeiten aktiv im Statistiklabor, lernen die statistischen Methoden aktiv in der direkten Anwendung.
- *Students learn to do well only what they practice doing*: Zusätzlich zu den Reports lösen die Studierenden wöchentlich Übungsaufgaben. Die statistischen Probleme können entweder manuell oder im Statistiklabor gelöst werden. Die Übungsaufgaben trainieren die Anwendung der in der Vorlesung vermittelten statistischen Methoden.
- *Learning is enhanced by having students become aware of and confront their misconceptions*: Die statistischen Reports werden jeweils in einer Übung von dem Dozenten diskutiert. Dabei werden Lösungsalternativen und ihre Vor- und Nachteile vorgestellt, Lösungen der Studierenden/Teams miteinander verglichen. Diese Vorgehensweise vermittelt verschiedene Herangehensweisen an statistische Probleme, mit denen sich die Studierenden bereits auseinandergesetzt haben. Die Vermittlung umfasst auch die Demonstration verschiedener Kombinationen statistischer Methoden.

Es wurden Bewertungsmethoden benötigt, um den Grad der Zielerreichung der neuen Vermittlungsmethoden zu messen. Es musste überprüft werden, inwiefern die Lernmodule und das Einsatzkonzept den Zielen und konstruktivistischen Prinzipien gerecht wurden. Im Folgenden werden ausgewählte Bewertungsmethoden diskutiert, die in das Einsatzkonzept integriert wurden.

**Individual or Group Projects:** Diese Methode ist in der entwickelten Reportkomponente in *Statistik interaktiv komplett* enthalten. Studierende (individuell oder im Team) sollen ein Problem eingrenzen, verstehen und mit Hilfe statistischer Methoden lösen. Die Dokumentation dieser Schritte soll in Form eines statistischen Reports stattfinden. Die von Garfield genannten Aspekte werden

<sup>283</sup>[Garfield1999], S. 2ff.

<sup>284</sup>Vgl. [Garfield1995], S. 30f.

umgesetzt: "understanding of the problem, collect data, analyze data, interpretation of data analysis, effective communication in written report."<sup>285</sup> Die kritische Reflexion über die Erstellung der Reports fasst Garfield unter dem Aspekt *Portfolios of student work* auf. Dort soll von den Studierenden ein mini research paper und ein reflective journal über eine statistische Problemlösung angefertigt werden. In der von uns entwickelten Reportkomponente wurden beide Aspekte zusammen umgesetzt. Das research paper wurde in Form des Reports erstellt. Die Reflexion fand sowohl im Report als auch in den jeweiligen Reportbesprechungen statt. Die Studierenden wurden angehalten, den Lösungsweg begründet darzustellen, z.B. das Ausschließen alternativer Wege darzulegen und zu begründen.

**Concept maps:** Diese Methode soll dazu beitragen, dass Studierende Verbindungen zwischen verschiedenen statistischen Methoden verstehen und selbst herstellen. Betrachten wir die Gestaltung der Lernmodule, so ist die Idee von Concept maps -Zusammenhänge zwischen Methoden herzustellen- durch die Verlinkungen der verschiedenen statistischen Methoden ansatzweise umgesetzt: Theoriekomponenten sind verlinkt mit entsprechenden Laborseiten, aber auch mit Theorieseiten anderer statistischer Gebiete. Concept maps sind also in der Struktur der Lernsoftware enthalten. Im Einsatzkonzept wurde eine Erweiterung vorgenommen: Die Studierenden erstellten selbst concept maps über vermittelte statistische Inhalte. Dieses Vorgehen betrifft die Vermittlung und Bewertung. Die gemeinsame Erstellung von concept maps mit den Studierenden vermittelt Zusammenhänge der Methoden, es findet Wissenskonstruktion statt. Dieser Aspekt wurde in den Aufgaben der Lehrveranstaltung<sup>286</sup> aufgegriffen: *Sie haben in den Lernmodulen zur deskriptiven Statistik verschiedene Methoden zur Analyse von Daten kennen gelernt. Listen Sie diese auf und setzen Sie sie in einem nächsten Schritt in eine sinnvolle Verbindung zueinander, die alle Zusammenhänge der erlernten statistischen Methoden/Konzepte untereinander aufzeigt.*

**Critics of statistical ideas or issues in the news:** Garfield schlägt vor, die Studierenden mit aktuellen Zeitungsartikeln zu konfrontieren und durch gezielte Fragestellungen eine kritische Auseinandersetzung der Studierenden mit diesen zu fördern. Die Aktualität der Statistik wird deutlich, die Studierenden merken, wie sie statistische Methoden auf das tägliche Leben übertragen können. Kombiniert man diese Methode mit der des statistischen Report, kann man die Studierenden die kritische Auseinandersetzung mit dem Artikel in Form eines statistischen Reports dokumentieren lassen. Als mögliche Fragen schlägt Garfield vor:<sup>287</sup>

1. "What do you think is the purpose of the research study described in this article?"
2. What method or methods were used to answer the research question?
3. What questions would you like to ask the investigators in order to better understand the study?
4. Are there any aspects of the study that might make you question the conclusions presented in this article?"

Die Beantwortung solcher Fragen setzt statistisches Denken und kritische Reflexion voraus. Diese muss gemeinsam mit den Studierenden trainiert werden. Dieser Aspekt wurde im Rahmen der zweiten Durchführung der Evaluationsstudie umgesetzt. Im Einsatzkonzept Ib wurde dies implementiert. Artikel und entsprechende Fragestellungen wurden im Lernmanagementsystem bereitgestellt.

**Minute papers:** Auch diese Methode wurde im Einsatzkonzept aufgegriffen. Minute Papers haben die Aufgabe, das Verständnis der Studierenden für gerade vermittelte statistische Konzepte und Methoden zu überprüfen. Außerdem kann die Fähigkeit der präzisen statistischen Dokumentation und Kommunikation gefördert werden. Innerhalb der letzten Minuten einer Vorlesung wird der Studierende dazu aufgefordert, in einem minute paper zum Ausdruck zu bringen, was er glaubt, verstanden zu haben. Es kann ihm auch eine konkrete Frage gestellt werden. Auf diese Weise

<sup>285</sup>[Garfield1999], S. 4.

<sup>286</sup>In der 2. Durchführung der Evaluationsstudie behandelt.

<sup>287</sup>[Garfield1999], S. 6.

bekommt der Dozent direktes Feedback von den Studierenden, kann Kommunikations- und Verständnisschwierigkeiten erkennen und diesen gezielt entgegenwirken.

**Performance assessment of statistical problem solving:** Eine der Hauptschwierigkeiten bestand in der Bewertung des erzielten Lernerfolges beim Studierenden. Da die Reportkomponente für uns von gesteigertem Interesse war, wurde sie fest im Einsatzkonzept verankert und gezielt hinsichtlich ihres Erfolges überprüft. Ausschlaggebendes Bewertungskriterium war die Entwicklung der Leistung der Studierenden - hier entstand ein Messproblem. Olsen schlägt folgende Skala vor, mit dessen Hilfe das statistische Wissen der Studierenden sowie ihre Kommunikationsfähigkeiten bewertet und gemessen werden sollen.

"A score of 4 on statistical knowledge means that the student completely understands a problem's statistical components, synthesize a correct relationship among these components, uses appropriate and correctly executed statistical techniques, and provides reasonably correct answers."<sup>288</sup>

Betrachten wir anschließend die Skala für communication skills:

"A score of 4 in communication scale means that the student's explanation of what was done was clear, complete, organized, and correct, that appropriate assumptions and caveats are stated, diagrams or plots used are appropriate, and an appropriate and complete conclusion is stated."<sup>289</sup>

Es scheint angebracht, die vorgestellten Vorschläge kritisch zu betrachten, insbesondere wenn sie eingesetzt werden, um Leistungen allein in Zahlen zu messen. Kombiniert mit qualitativen Messmethoden können sie allerdings zu sinnvollen Messergebnissen führen.

Bezüglich des statistischen Report ist eine gemeinsame Anwendung beider Skalen positiv zu bewerten, da die Kommunikation ein wichtiger Bestandteil des Verfassens statistisches Reports ist. Wir schlagen im Rahmen dieser Arbeit eine Ergänzung vor. Zusätzlich zu der quantitativen Messung den beiden Skalen gemäß muss der Versuch einer qualitativen Messung unternommen werden. So muss der Entwicklungsprozess des Studierenden herangezogen werden. Seine Entwicklung beim Verfassen der Reports sollte in die Bewertung eingehen. Auf diese Weise kann der Beitrag des konstruktivistisch gestalteten Einsatzkonzeptes und der Lernsoftware festgestellt werden. Im Einsatzkonzept wurde der Versuch unternommen, ein Bewertungsschema für die statistischen Reports festzulegen. Dieses sieht die Messung nach qualitativen Kriterien vor.

**Bewertungsschema für statistische Reports** Durch die Reports wurde gemessen, ob der Studierende statistisches Problemlösen während des Semesters erlernte. Dies konnte durch quantitative Evaluationsmethoden nicht untersucht werden. Da die Bewertung statistischer Reports aufwendig ist, wurde ein Schema entworfen: Zum einen vermindert dies den Bewertungsaufwand des Dozenten, zum anderen findet eine Standardisierung der Bewertung statt. Das Schema enthält die folgenden Aspekte:

- Wurde der Report strukturiert in Problemstellung, Methodenauswahl, Interpretation, Rückschluss zur Realität?
- Fand eine begründete Auswahl der statistischen Methoden statt?
- Wurde die Reihenfolge der Methodenanwendungen begründet?
- Wurden Zusammenhänge zwischen den Methoden und erzielten Ergebnissen hergestellt?
- Wurden die Methoden korrekt angewendet und umgesetzt?
- Wurden die erzielten analytischen Ergebnisse statistisch und inhaltlich interpretiert?

---

<sup>288</sup>[Garfield1999], S. 7.

<sup>289</sup>[Garfield1999], S. 7.

Es empfiehlt sich, aufgrund der genannten Vorteile, die Bewertung statistischer Reports bei zukünftigen Einsätzen einem Bewertungsschema gemäß vorzunehmen.

**Zwischenfazit:** Die Betrachtungen einiger von Garfield und Gal vorgeschlagenen neuen Bewertungs- und Vermittlungsmethoden wurden auf das Einsatzkonzept und somit auf die Statistikausbildung übertragen. Es wurde gezeigt, dass neue Bewertungs- und Vermittlungsmethoden zu einer Neugestaltung der Statistikausbildung führen können. Dabei können die neuen Medien eine wichtige Rolle spielen, sofern sie zielgerichtet eingesetzt werden. Im Rahmen der Evaluationsstudie wurde die Qualität der Integration einiger der neuen Bewertungsmethoden in die Einsatzkonzepte und somit in die Statistikausbildung überprüft.

### 11.3 Qualitative und quantitative Evaluation

In der Evaluationsstudie wurde zwischen *qualitativer* und *quantitativer* Evaluation unterschieden. Es handelt sich um verschiedene Evaluationsmethoden, die im Folgenden für diese Arbeit definiert und gegeneinander abgegrenzt werden.

Ein mögliches Abgrenzungskriterium ist die Ebene der Betrachtung:

- Qualitative Evaluation betrifft vorwiegend das Individuum.
- Quantitative Evaluation betrifft vorwiegend das Kollektiv, die Allgemeinheit, die Teilnehmer als Gruppe betrachtet. Quantitative Evaluationsverfahren ermöglichen Aussagen über Anteile der betrachteten Stichprobe.<sup>290</sup>

Hier wird von *qualitativer Evaluation* gesprochen, wenn der Einsatz multimedialer Lernsoftware mit Hilfe qualitativer Methoden evaluiert wird, z.B. durch die Methode des teilnehmenden Beobachtens, durch offene Befragungen der Studierenden, durch Minute Papers, durch Concept Maps, durch die statistischen Reports und das Abschlussgespräch mit den Teams. Die von Garfield vorgeschlagenen neuen Vermittlungs- und Bewertungsmethoden sind hier anzusiedeln, da sie die Betrachtung des Individuums ermöglichen.

Bei der *quantitativen Evaluation* werden hier statistische Analysemethoden auf die Daten der Gruppe angewendet. Die Studierenden werden schriftlich befragt. Die Antwortmöglichkeiten sind so konzipiert, dass sie mit Hilfe statistischer Verfahren ausgewertet werden können. Das Vorhandensein bestimmter Eigenschaften in der Gruppe der Teilnehmer wird evaluiert.

Das Ziel bestand darin, eine Kombination qualitativer und quantitativer Evaluationsmethoden zu entwickeln, die flächendeckend umsetzbar ist und die jeweiligen Vorteile beider Methoden nutzt. In der Evaluationsstudie wurde zunächst qualitativ, dann quantitativ, und schließlich kombiniert evaluiert. Dieses Vorgehen resultierte aus den Erfahrungen, die im Verlauf der Evaluationsstudie gewonnen wurden.

### 11.4 Varianten Ia und Ib des Einsatzkonzeptes

Nach diesen ersten Einordnungen und Abgrenzungen findet eine Gegenüberstellung der Varianten Ia und Ib des schrittweise entwickelten Einsatzkonzeptes statt. In der explorativen Evaluationsstudie wurden beide Varianten entwickelt und evaluiert. Beide organisieren den Einsatz der multimedialen Lernsoftware *Statistik interaktiv komplett* im Rahmen einer Universitätsveranstaltung. Durch schrittweise Verbesserung entstand *ein* Einsatzkonzept für das gewählte Einsatzszenario. Die in der 1. Durchführung gewonnenen Erfahrungen, die festgestellten Vor- und Nachteile bestimmter Komponenten des Einsatzkonzeptes sowie der Lernsoftware wurden beim Aufstellen des Einsatzkonzeptes Ib berücksichtigt. Die 1. Durchführung der Evaluationsstudie hatte Pilotprojektcharakter. Sie diente der Sammlung erster Erfahrungen und Erkenntnisse über den Einsatz multimedialer

<sup>290</sup> Allerdings ist hier auch die eine Betrachtung einzelner Individuen möglich, z.B. durch die Identifikation von Ausreißern.



Lernsoftware in der statistischen Grundausbildung, der Aufdeckung von Schwächen des Einsatzkonzeptes, der Lernsoftware und ihrer Module/Komponenten. Dies geschah durch teilnehmendes Beobachten. Diese Tätigkeiten wurden unter dem Begriff „explorative“ Evaluation subsumiert. Die gezielte Kombination verschiedener Komponenten der Lernsoftware und des Einsatzkonzeptes zur Erreichung bestimmter Ziele fand erst in der 2. Durchführung statt.

Gemeinsamkeiten und Unterschiede der Varianten des Einsatzkonzeptes werden im Folgenden komprimiert dargestellt.<sup>291</sup>

**Gemeinsamkeiten** der Varianten des Einsatzkonzeptes:

- Beide organisieren den Einsatz der multimedialen Lernsoftware *Statistik interaktiv komplett* in einer Lehrveranstaltung an der Universität, betrachten dasselbe Einsatzszenario.
- Beide haben die übergeordneten Ziele, die Statistikausbildung zu verbessern, speziell: Defizite der Statistikausbildung entgegenzuwirken.
- Beide sehen eine Betreuung der Teilnehmer der multimedialen Veranstaltung vor, jedoch in verschiedener Intensität.
- Beide enthalten die Komponenten Vorlesung, Übung/Tutorium und betreute Rechnerstunde.
- In beiden kommen dem Statistiklabor und dem Vermitteln des statistischen Reportverfassens eine zentrale Rolle zu.

Die **Unterschiede** liegen im Detail:

- Während das Einsatzkonzept Ia eine Unterteilung der Vorlesung in eine Motivations- und eine Theorievorlesung vorsieht, findet auf Basis des Einsatzkonzeptes Ib eine Vorlesung statt, die sowohl Motivations- als auch Theorieteile enthält.
- Das im Einsatzkonzept Ia enthaltene Tutorium wird im Einsatzkonzept Ib durch eine Übung ersetzt. Der Unterschied besteht darin, dass die Übung von einem Diplomanden geleitet wird, der im Rahmen des Forschungsprojektes beschäftigt ist. Die Tutorien werden von zwei studentischen Hilfskräften geleitet, die sich am Anfang ihres Hauptstudiums befinden. Die Gruppen wurden geteilt, jeder Tutor betreut eine Gruppe. Im Gegensatz dazu findet die Übung mit allen Teilnehmern in einer Gruppe statt.
- Eine betreute Rechnerstunde wird im Einsatzkonzept Ib von dem Diplomanden geleitet, während es im Einsatzkonzept Ia einen gesonderten Tutor gibt, der diese Aufgabe übernimmt, mit den eigentlichen Tutorien aber in keiner Verbindung steht.
- Im Einsatzkonzept Ia werden die Motivations- und Theorievorlesung von zwei verschiedenen Dozenten gelesen, während die Vorlesung im Einsatzkonzept Ib durchgängig von einem Dozenten gehalten wird.
- Die im Einsatzkonzept Ia etablierte Prüfungsform sieht zusätzlich zu den während des Semesters anzufertigenden statistischen Reports eine Multiplechoiceklausur vor, auf die beim Einsatzkonzept Ib verzichtet wird. Dieses schließt die Prüfungsleistung dafür mit einer mündlichen Prüfung jedes Teams ab.
- Die im Einsatzkonzept Ia eingesetzten Komponenten von *Statistik interaktiv komplett* wurden weiterentwickelt und in ihrer neuen Version im Einsatzkonzept Ib eingesetzt. Insbesondere das Statistiklabor wurde auf die im ersten Einsatz erkannten Bedürfnisse der Studierenden/Dozenten abgestimmt. Eine Berichterstellungsfunktion wurde entwickelt, die das Verfassen und Versenden statistischer Reports organisiert, ermöglicht und erleichtert.
- Beide Varianten Ia und Ib des Einsatzkonzeptes legen einen Schwerpunkt auf die Reporterstellung. Der entscheidende Unterschied besteht darin, dass nach den Erfahrungen der 1. Durchführung neue reportspezifische Komponenten entwickelt, eingesetzt und evaluiert wurden: ein Musterreport, eine Coaching-Komponente, eine Animation zur schrittweisen Reporterstellung.

---

<sup>291</sup>Die folgenden Ausführungen verdeutlichen die Unterschiede und Gemeinsamkeiten an konkreten Beispielen.

- Der Einsatz der multimedialen Lernsoftware *Statistik interaktiv komplett* wird im Einsatzkonzept Ib in das Lernmanagementsystem Blackboard eingebettet, während der Austausch von Materialien im Einsatzkonzept Ia über ein Kurslaufwerk im Netzwerk (Novell) der Universität stattfindet.
- Das Einsatzkonzept Ia unterteilt die Veranstaltung in zwei Phasen. Dabei wird das statistische Report verfassen erst in der zweiten vermittelt, nachdem zunächst der Umgang mit dem Statistiklabor gefestigt wurde. Im Einsatzkonzept Ib erfolgen diese beiden Schritte parallel zueinander, d.h. der Studierende wird über das Verfassen statistischer (Mini-)Reports an das Statistiklabor herangeführt.

Die Darstellung der Gemeinsamkeiten und Unterschiede der Varianten des entwickelten Einsatzkonzeptes verdeutlicht, dass trotz großer Ähnlichkeiten entscheidende Unterschiede im Detail bestehen. Es ist anzumerken, dass viele Komponenten des Einsatzkonzeptes Ia erst während der 1. Durchführung herausgebildet wurden, da eine zu geringe Vorlaufzeit bestand. Das Forschungsprojekt lief zu spät und zu langsam an, um den gesamten Einsatz vor Beginn des Semesters vollständig zu konzipieren. Um die Qualität der entwickelten Lernmodule möglichst früh zu überprüfen, mussten diese ohne großen zeitlichen Vorlauf in der Lehre eingesetzt werden. Die Projektgruppe Bielefeld nahm sich dieser Aufgabe an und musste den Projektpartnern die Notwendigkeit der Entwicklung eines Einsatzkonzeptes erst vor Augen führen. Während der ersten Durchführung und der schrittweisen Entwicklung des Einsatzkonzeptes wurde deutlich, dass erkannte Schwächen und Stärken einzelner Komponenten in den Entwicklungsprozess der Lernsoftware und des Einsatzkonzeptes integriert werden mussten. Dies betraf insbesondere die Stabilität und Funktionalität des Statistiklabors, wie z.B. die Notwendigkeit der Entwicklung und Implementation der Berichterstellungsfunktion.

Die Durchführung der Evaluationsstudie folgte keiner festen Reihenfolge, wie z.B. der folgenden:

1. Entwicklung eines Einsatzkonzeptes
2. Evaluation des Einsatzkonzeptes
3. Ableitung von Änderungen des Einsatzkonzeptes aufgrund der Evaluationsergebnisse
4. Entwicklung des zweiten Einsatzkonzeptes
5. Evaluation des zweiten Einsatzkonzeptes

Die Schritte Entwicklung und Evaluation erfolgten bei der 1. Durchführung aufgrund der geringen Vorlaufzeit und des nicht organisierten Projektablaufs des Forschungsprojektes *Neue Statistik* fast zeitgleich. Während der Durchführung wurden Veränderungen des Einsatzes vorgenommen, wenn die Notwendigkeit dazu bestand, z.B. wurde das Statistiklabor durch die Entwicklung entsprechender Bibliotheken intensiver genutzt. Die aus den Evaluationsergebnissen resultierenden Konsequenzen wurden in der 2. Durchführung berücksichtigt. Hier fand eine direkte Verbindung statt zwischen:<sup>292</sup>

- dem aufgestellten Defizitmodell,
- den Komponenten der Lernsoftware und
- den Komponenten des Einsatzkonzeptes

Nachdem mögliche Kombinationen verschiedener Komponenten in der 1. Durchführung der Evaluationsstudie experimentell erforscht wurden, wurden in der 2. Durchführung konkrete Kombinationen und damit verbundene Hypothesen evaluiert.

---

<sup>292</sup>Vgl. 13.1.3.

## 12 Explorative Evaluationsstudie - 1. Durchführung

In diesem Kapitel wird über die erste Durchführung der Evaluationsstudie reflektiert. Dem Leser werden Schlüsselerlebnisse und -erfahrungen beschrieben. Ausgewählte Komponenten des Einsatzkonzeptes Ia werden dargestellt. Es handelt sich nicht um einen alle Aspekte der 1. Durchführung abdeckenden Bericht. Dem Leser werden bezüglich ausgewählter Komponenten des Einsatzkonzeptes zwei Aspekte gemeinsam dargestellt:

- eine Beschreibung der Komponenten, ihrer Kombinationen und der dahinterstehenden Ideen,
- ein Evaluationsbericht, der dem Leser Einblicke in die mit der 1. Durchführung entstandenen Probleme gibt, interessante Beobachtungen schildert und erste Evaluationsergebnisse herausstellt.

Die Darstellung der ersten Durchführung erfolgt in Form kleiner „stories“, erscheint dem Autor angemessen und zugleich interessant zu sein, da sie Pilotcharakter besaß und zum Sammeln von Erfahrungen in unterschiedlichen Bereichen diente. Welche Überlegungen mit der Konzeption und Durchführung einer Evaluationsstudie verbunden sind, wird strukturiert im Rahmen der 2. Durchführung dargestellt.

Nach einem kurzen Überblick über Rahmenbedingungen und Startschwierigkeiten der 1. Durchführung werden dem Leser vier „stories“ zu den folgenden Themen beschrieben:

- Statistische Motivations- und Theorievorlesung
- Statistische Reports
- Evaluationsinstrument Tutor - Tutorenmanagement
- Neue Prüfungsformen

Der Autor wird als Projektteilnehmer und teilnehmender Beobachter über die Erfahrungen und Erkenntnisse reflektieren.

### 12.1 Rahmenbedingungen und Startschwierigkeiten

An der Universität Bielefeld wurde im Wintersemester 01/02 die obligatorische Statistik I Vorlesung des Grundstudiums in den Diplomstudiengängen Volkswirtschaftslehre und Betriebswirtschaftslehre auch als multimediale Veranstaltung angeboten. Dies erfolgte über die Einrichtung eines Wahlpflichtfaches *Statistik interaktiv!*, dessen Teilnehmerzahl aus technischen Gründen auf 32 begrenzt war. Das Projekt wurde in der ersten Woche des Semesters in der Orientierungsphase für Erstsemester durch die Projektleitung vorgestellt. Anschließend lagen Listen aus, auf denen sich Interessierte eintragen und für das Projekt anmelden konnten. Die Auswahl der Teilnehmer aus den Studierenden, die sich für das Projekt anmeldeten, erfolgte nach dem Zufallsprinzip.

Bei der Auswahl der Teilnehmer traten erste Schwierigkeiten auf. Die Teamer, Grund- oder Hauptstudiumsstudierende, die die Erstsemester durch die Orientierungsphase begleiteten, waren nicht ausreichend über das Projekt informiert, konnten somit die von den Erstsemestern gestellten Fragen zum Projekt nicht hinreichend beantworten. Es scheint nicht ausreichend zu sein, die Teamer mündlich von dem Projekt zu unterrichten und ihnen Anmelde Listen zu übergeben. Bei der 2. Durchführung wurden deshalb die wichtigsten organisatorischen sowie inhaltlichen Aspekte bezüglich des Projektes schriftlich an die Teamer verteilt. Es wurde auf Ansprechpartner sowohl für die Teilnehmer als auch für Erstsemester verwiesen.

Von Semesterbeginn an hatte jeder Teilnehmer die Möglichkeit, sich im Hochschulrechenzentrum mit der Multimediasoftware vertraut zu machen. In der ersten Veranstaltung wurden Benutzerausweise an die Teilnehmer verteilt und Prioritätszeiten an den Rechnern der Universität eingeräumt, was mit einem hohen organisatorischen Aufwand verbunden war. Es stellte sich heraus, dass der Einsatz von Multimedia grundsätzlich mit einem Mehraufwand betrieben werden musste. Dass der Aufwand nicht nur auf der Seite der Dozenten lag, zeigt die Aussage eines Studierenden: „Wenn man nun letztendlich den wöchentlichen Zeitaufwand für Statistik betrachtet, kommt die Frage auf, ob man BWL oder Statistik studiert!“

Den Teilnehmern wurde in der dritten Vorlesungswoche die Multimediastsoftware kostenlos von der Projektleitung auf CD-Rom zur Verfügung gestellt. Auffallend war, dass einige Teilnehmer Schwierigkeiten mit der Installation des Programms hatten, da die CD keine Autostartfunktion besaß. Dies sollte bei zukünftigen Einsätzen beachtet werden. Die Studierenden wurden bei der 1. Durchführung nur wenig betreut, was die Einarbeitung in die Komponenten der Multimediastsoftware betrifft. Es wurde untersucht, inwieweit die Studierenden eigenständig lernen können. Wir sprechen in diesem Zusammenhang vom *self-paced-learning*. In der 2. Durchführung wurden sie stärker betreut. Es gab mehrere Einführungen seitens der Projektleitung, die den Studierenden sowohl technische als auch inhaltliche Aspekte der Software näher brachten. Diese Betreuung bezeichnen wir als *guided learning*.

Von den 60 in der Anmeldeliste Eingetragenen wurden nach dem Zufallsprinzip 32 Teilnehmer ausgewählt.

Aufgrund der Tatsache, dass die Anmeldung zu dem Projekt an keine Qualifikationen der Teilnehmer gekoppelt war, entstand eine in sich sehr heterogene Gruppe. Viele Teilnehmer waren mit dem Umgang mit dem Rechner vertraut, hatten zum Beispiel keine Probleme mit der Installation, dem Umgang mit dem Labor. Andere jedoch hatten große besonders technische Schwierigkeiten, die sie beim Lösen statistischer Probleme mit Hilfe des Rechners sehr viel Zeit kosteten. Dies konnte teilweise auf mangelnden Einsatz aber auch auf zu geringes technisches Vorwissen der Teilnehmer zurückgeführt werden.

*Organisatorische und technische Voraussetzungen:* Sowohl die beiden Vorlesungen, als auch die Tutorien fanden in Räumen statt, die mit einem Beamer ausgestattet waren. So konnte jederzeit auf die Lernsoftware zugegriffen werden. Außerhalb der betreuten Rechnerstunden wurden den Projektteilnehmern Prioritätszeiten beim HRZ (Hochschulrechenzentrum) eingeräumt. Dies hatte zum Ziel, den Projektteilnehmern viele Möglichkeiten zu bieten, sich mit der Lernsoftware auseinanderzusetzen. Ihre Lernbereitschaft/Motivation sollte nicht dadurch gebremst werden, dass ihnen kein Rechner zur Verfügung stand.

## 12.2 Ausgewählte Komponenten des Einsatzkonzeptes Ia

Im Folgenden werden ausgewählte Komponenten der 1. Durchführung der Evaluationsstudie skizziert, Erfahrungen diskutiert. Die Entwicklung jeder Komponente des Einsatzkonzeptes war mit zwei Aspekten verbunden:

- Einem Gestaltungsproblem: Was will man mit der Komponente erreichen? Wie erreicht man es?
- Einem Messproblem: Was wurde mit den Komponenten erreicht?

### 12.2.1 Statistische Motivations- und Theorievorlesung

Unterteilt man die Veranstaltung in zwei aufeinander folgende Phasen, so bestand der Schwerpunkt der ersten Phase darin, die Studierenden mit dem in der Software enthaltenen Statistiklabor vertraut zu machen. Die zweite Phase der Veranstaltung konzentrierte sich darauf, die Studierenden in die Erstellung statistischer Reports einzuführen. Dabei stand das Verfassen strukturierter Texte und deren Kombination mit den im Statistiklabor erarbeiteten Ergebnissen im Mittelpunkt. Dem Studierenden sollte vermittelt werden, dass die Dokumentation einen wichtigen Bestandteil statistischen Problemlösens darstellt.

Pilotprojekte<sup>293</sup> hatten gezeigt, dass den Studierenden das Schreiben von Texten erst vermittelt werden muss. Zudem waren sie überfordert, wenn ihnen statistische Probleme, die Theorie der statistischen Methoden und der Umgang mit dem Statistiklabor gleichzeitig vermittelt wurden. Als

<sup>293</sup>Statistikveranstaltungen, in denen vor der Evaluationsstudie Teile der Vorgängersoftware *Statistik interaktiv!* eingesetzt wurden.

Konsequenz fand im Einsatzkonzept Ia eine Differenzierung zwischen vier Veranstaltungskomponenten statt. „Überall ein wenig macht ein Ganzes!“, kommentierte ein Teilnehmer die Organisation der Veranstaltung.

1. Motivationsvorlesung: Statistische Probleme wurden insbesondere durch den Einsatz des Statistiklabors demonstriert. Das Ziel bestand in der Vermittlung der Ideen statistischer Methoden. Es wurde auf verschiedenen Laborseiten demonstriert, wie diese auf reale Datensätze angewendet werden könnten, um Probleme zu lösen.
2. Theorievorlesung: Die theoretischen Hintergründe der in der Motivationsvorlesung eingesetzten Methoden wurden vermittelt. Dies erfolgte ohne den Einsatz des Statistiklabors.
3. Tutorium: Der Umgang mit den statistischen Methoden wurde in Gruppen aufgabenbezogen trainiert. Der Tutor trat als Guide/Moderator im Sinne des Konstruktivismus auf.
4. Betreute Rechnerstunde: Statistische Fragestellungen wurden mit Hilfe des Statistiklabors gelöst. Es stand ein Tutor zur Beantwortung von Fragen bereit, nicht zur Demonstration von Lösungen. Die durch die Tutoren betreuten Rechnerstunden erfüllten den Zweck, dass die Studierenden während des Lösens der Übungsaufgaben mit Hilfe der multimedialen Lernsoftware einen Ansprechpartner bei auftretenden technischen oder inhaltlichen Problemen hatten. Der Tutor sollte vor Ort sein, um in Notfällen einzugreifen. Er sollte sich allerdings bei Fragen inhaltlicher Art mit seinen Hilfestellungen zurückhalten, um eine anschließende Bewertung der Eigenleistung der Studierenden gewährleisten zu können. Die Studierenden sollten durch eigene Aktivität lernen, wie Statistik von einem *Programm* umgesetzt werden kann.

**Reflexion:** Generell äußerten sich die Teilnehmer positiv über die Veranstaltung. Die Unterteilung der Vorlesung in Motivations- und Theorievorlesung wurde positiv bewertet: „Die Unterteilung ist sinnvoll, da man somit einen Plan hat wie alles läuft, d.h. man hat eine festgelegte Route, der man folgt, um ans Ziel zu kommen.“ Die Teilnehmer fühlten sich in der Lage, sich zunächst (Motivationsvorlesung) darauf einzulassen, nur die Ideen der Methoden zu verstehen, weil sie sich bewusst waren, dass später (Theorievorlesung) alle Aspekte geordnet wieder aufgegriffen und detailliert erklärt wurden. Sie ließen sich darauf ein, zu verstehen, dass die Statistik in verschiedenen Bereichen auftaucht und statistische Kenntnisse deshalb wichtig sind. Besonders der Realitätsbezug der Veranstaltung wurde von den Teilnehmern als positiv empfunden. Einige Studierende merkten allerdings an, dass die Trennung der Vorlesungen bei ihnen dazu führte, nur die Theorievorlesung zu besuchen. - „Das Problem hierbei ist wohl, dass man den Stoff auch bewältigen kann, wenn man die Motivationsvorlesung auslässt.“ Ein Großteil betonte, dass sie viel motivierter als in anderen Fächern waren, die Methoden zu erlernen, weil sie immer wieder auf reale Fragestellungen bezogen wurden, in denen ihnen die statistischen Kenntnisse weiterhelfen können. Viele Teilnehmer setzten sich regelmäßig (ungefragt) mit der Projektleitung in Verbindung und berichteten von Erlebnissen, bei denen sie ihre bisher erlernten Kenntnisse zum besseren Verständnis anwenden konnten.<sup>294</sup>

Das schrittweise Heranführen an das Statistiklabor hatte positive Effekte. Die Demonstration von Laborseiten in der Motivationsvorlesung weckte Interesse und zeigte Möglichkeiten, den Rechner und die Statistik zu kombinieren. Die Präsentation individuell erstellter Laborseiten führte zu Sicherheit im Umgang mit den Laborobjekten. Die Betreuung der Teilnehmer in der Rechnerstunde war anfangs stark aufgrund technischer Startschwierigkeiten. Im Laufe des Semesters wurde kontinuierlich weniger Betreuung der Arbeit im Statistiklabor nachgefragt.

In der **Theorievorlesung** wurde das Statistiklabor nur eingesetzt, wenn an Fragestellungen aus der Motivationsvorlesung direkt angeknüpft wurde. Die entsprechenden Laborseiten wurden aufgerufen. Es fand keine aktive Arbeit im Statistiklabor statt, um eine reine Vermittlung der Theorie der Methoden vorzunehmen. Die 1. Durchführung zeigte, dass die Trennung von Theorie und Anwendung problematisch war. Als Konsequenz wurden Lösungsansätze entwickelt, die Theorie und

<sup>294</sup>Einige berichteten z.B. über die grafische Aufbereitung von Wahlergebnissen in Zeitungen, kritisierten verzerrende Darstellungen.

Anwendung kombinieren, wie dies in den in Teil I vorgestellten Umsetzungsbeispielen der Fall ist. Zu Beginn der Theorievorlesung wurde den Studierenden geraten, diese als Hilfe zur Erstellung einer individuellen statistischen Formelsammlung anzusehen. Wichtige Formeln oder Definitionen wurden in der Vorlesung hervorgehoben, um ihre Relevanz für die Studierenden zu unterstreichen. Die Studierenden wurden darauf hingewiesen, dass von ihnen erwartet wird, sich die Formelsammlung selbstständig nach dem konstruktivistischen Prinzip des *self-paced-learning* zu erstellen. Die abgegebenen Formelsammlungen ließen vereinzelt auf Lücken schließen, deuteten aber zum Großteil auf eine ernsthafte Auseinandersetzung mit dem gelehrteten Stoff hin.

Zusammenfassend konnten die Unterteilung in verschiedene Veranstaltungskomponenten und die eigenständige Erstellung der Formelsammlung positiv bewertet werden. Die Formelsammlung folgte dem konstruktivistischen Prinzip, sich *self-paced* Wissen anzueignen. In der 2. Durchführung der Evaluationsstudie wurden diese positiv bewerteten Komponenten nicht wiederholt evaluiert. Dies ist auf die Vielzahl der zu untersuchenden Komponenten zurückzuführen.

### 12.2.2 Statistische Reports

Die zentrale Komponente des Einsatzkonzeptes war das Erstellen statistischer Reports. Diese steht in enger Beziehung zu der Motivationsvorlesung, die das Ziel verfolgte, den Studierenden statistische Problemlösung zu demonstrieren. Statistische Verfahren wurden dann vermittelt, wenn sie zur Lösung eines Problems benötigt wurden. Der statistische Report wurde als Vermittlungs- und Messinstrument eingesetzt. Der Studierende wurde trainiert, Probleme strukturiert zu lösen. Es wurde überprüft, ob der Studierende in der Lage war, die Zusammenhänge zwischen den vermittelten Inhalten beim statistischen Problemlösen zu einem Netz sinnvoll zusammenzuknüpfen. Die in der Evaluationsstudie entwickelte Prüfungsform sieht vor, dass die Studierenden im Laufe des Semesters drei statistische Reports verfassen. Es bedurfte der Entwicklung einer Prüfungsform, die weit über das Abfragen statistischer Methodenkompetenzen hinausgeht.

Die Reports wurden inhaltlich in das Vorlesungsschema eingegliedert. Auf der einen Seite sollte das Verständnis für statistische Methoden, die in der Vorlesung vermittelt wurden, durch die Reports überprüft werden. Zum anderen dienten die Reports dazu, zu überprüfen, inwieweit die Studierenden in der Lage waren, sich statistische Methoden zur Lösung der gestellten Aufgabe anzueignen. Die Studierenden setzten sich auf diese Weise selbstständig mit statistischen Methoden auseinander. Hiermit wird das konstruktivistische Prinzip des *self-paced-learning*s verfolgt. Im Einsatzkonzept Ib wurden die Studierenden aufgrund der gesammelten Erfahrungen beim Verfassen statistischer Reports stärker betreut. Dies folgte dem konstruktivistischen Prinzip des *guided-learning*s. Dem Leser werden die Ergebnisse der teilnehmenden Beobachtung in Bezug auf die statistischen Reports im Folgenden komprimiert dargestellt.

**Beobachtungen zum 1. Report** Den Studierenden wurde ein Ausschnitt eines größeren Datenbestands bereitgestellt. Es folgte eine bewusst unpräzise formulierte Aufgabenstellung: *Führen Sie eine Datenanalyse mit den Ihnen aus der Vorlesung bekannten statistischen Methoden durch. Schreiben Sie einen statistischen Report, der Ihre Erkenntnisse und Ihre Vorgehensweise beschreibt.*<sup>295</sup>

Es sollte untersucht werden, inwiefern die Studierenden ohne klare Aufgabenstellung in der Lage waren, statistische Methoden zu selektieren und gezielt miteinander zu kombinieren.

Die von den Teilnehmern abgegebenen statistischen Reports spiegelten grundsätzlich hohe Lernerfolge wider. Die Teilnehmer gingen motiviert an die Erstellung des Reports heran, gaben sich Mühe, viele Aspekte zu berücksichtigen, die sie in der Veranstaltung erlernt hatten. Es wurde differenziert zwischen für die Problemlösung geeigneten und eher ungeeigneten statistischen Methoden. Keiner der abgegebenen Reports ließ Rückschlüsse auf große statistische Verständnisprobleme zu. Auffällig war, dass die Teilnehmer größtenteils das Statistiklabor und MS Excel zur Lösung einsetzten. Die Tutoren berichteten dies auch über die wöchentlich abzugebenen Übungsaufgaben. Eine Be-

<sup>295</sup>Die Problemstellung zur Erstellung des ersten Reports wurde zu Beginn der fünften Vorlesungswoche, an die Teilnehmer übergeben. Die Bearbeitungszeit betrug drei Tage.

fragung der Teilnehmer ergab, dass sie MS Excel präferierten, weil es in der Anwendung weniger Schwierigkeiten bereitete als das Labor. Es wurde bewusst darauf verzichtet, die Teilnehmer dazu zu zwingen, ausschließlich das Labor zu benutzen, um solche Aspekte herauszufinden.

Einige der Studierenden erstellten den ersten Report manuell, ohne Einsatz des Statistiklabors. Der Umfang des Datenmaterials ließ dies zu. Daraus wurde die Konsequenz abgeleitet, beim zweiten statistischen Report Daten vorzugeben, die eine manuelle Bearbeitung durch ihren Umfang verhinderten.

**Beobachtungen zum 2. Report** Im Vergleich zum 1. Report waren Fortschritte der Teilnehmer deutlich erkennbar. Sie verbesserten sich in ihrer Argumentation sowie der gezielten Auswahl und Anwendung statistischer Methoden.<sup>296</sup> Die Studierenden sollten in diesem Report zwei Datensätze analysieren. Auch hier wurde die Aufgabenstellung bewusst unpräzise formuliert, da die Studierenden im späteren Berufsleben oft vor dem Problem stehen werden, nicht klar formulierte Aufgaben mit Hilfe geeigneter Methoden lösen zu müssen. Es sollte zudem überprüft werden, ob die Studierenden in der Lage sind, die Entscheidung zu treffen, welche statistischen Methoden geeignet sind, um das gestellte Probleme zu lösen. Klare Aufgabenstellungen würden Methoden vorgeben und diesen Prozess verhindern.

Die Studierenden berichteten, dass ihre Probleme weniger im Technischen bestanden als vielmehr in der Interpretation der Ergebnisse. Sie probierten viele Methoden aus, versuchten die Daten geeignet zu analysieren, konnten aber keine zufrieden stellenden Aussagen über die Daten treffen. Sowohl die Reports als auch Gespräche mit einzelnen Gruppen machten das große Engagement der Teilnehmer und ihr Interesse für die Statistik deutlich. Die Schwäche des 2. Reports lag eindeutig in der zu umfangreichen Problemstellung. Dies wurde bei dem 3. Report berücksichtigt.

**Beobachtungen zum 3. Report** Bei einem Großteil der Studierenden war eine starke positive Entwicklung erkennbar. Sie gingen strukturiert an die Lösung des Problems heran, beschrieben und begründeten ihre Vorgehensweise.<sup>297</sup> Die Interpretation der erzielten Analyseergebnisse ließ erkennen, dass die statistischen Methoden verstanden wurden. Die Interpretation wurde als ein wichtiger Bestandteil statistischen Problemlösens angesehen. Die von den Studierenden angefertigten Reports zeigten, dass viel Zeit investiert wurde, dass außerdem über die anzuwendenden statistische Methoden nachgedacht wurde und nicht mehr nur die zuletzt gelernten Methoden unüberlegt angewendet wurden. Es wurden Methoden angewendet, deren Vermittlung teilweise weit zurück lag. Dies zeigte, dass im Verlauf des Semesters statistische Problemlösungskompetenz vermittelt wurde.

**Fazit zu den Reports** Grundsätzlich konnte bezüglich des Verfassens statistischer Reports ein äußerst positives Fazit gezogen werden. Jeder Projektteilnehmer gewann Erkenntnisse bezüglich der Lösung statistischer Probleme und der damit verbundenen Erstellung statistischer Reports. Da die Besprechung der statistischen Reports durch die Projektleitung positiven Einfluss auf die Qualität der im Anschluss angefertigten Reports zu haben schien, wurden die Reportbesprechungen in der 2. Durchführung stark intensiviert.

Dieser Übergang vom Prinzip des *self-paced-learning*s zum *guided-learning* zieht sich durch die gesamte Betreuung der Studierenden in der 2. Durchführung der Evaluationsstudie, insbesondere in Bezug auf den Prozess des Verfassens statistischer Reports. Die Anfertigung der statistischen Reports während des Semesters wurde als positiver Druck empfunden, sich ernsthaft mit den statistischen Methoden auseinanderzusetzen. „Die Idee, die Prüfungsleistungen schon während des Semesters zu erbringen, fand ich klasse. Man ist sozusagen gezwungen, kontinuierlich mitzuarbeiten, und kann somit den Faden nicht verlieren.“ (Zitat eines Teilnehmers)

Die identifizierten Schwächen und Instabilitäten des Statistiklabors wurden in der sich anschließenden Entwicklungsphase behoben. Das Statistiklabor wurde deutlich verbessert.

---

<sup>296</sup>Der zweite Report wurde zu Beginn der achten Vorlesungswoche ausgeteilt. Die Bearbeitungszeit betrug vier Tage, weil die Problemstellung umfangreicher als die des 1. Reports war.

<sup>297</sup>Der dritte statistische Report wurde Ende Januar gestellt und vier Tage lang von den Studierenden bearbeitet.

### 12.2.3 Evaluationsinstrument Tutor - Tutorenmanagement

Die Betreuung, Ausbildung und Beobachtung der Teilnehmer fand in der 1. Durchführung durch zwei Dozenten und drei Tutoren statt. Zwei Tutoren betreuten die Übungen, der dritte stand für technische Fragen in und außerhalb der Rechnerstunde zur Verfügung. Den Tutoren kam in der ersten Durchführung eine besondere Rolle zu. Sie wurden als Evaluationsinstrument eingesetzt. Statt wie im klassischen Sinn die Übungen zu leiten, bestand eine ihrer Zusatzaufgaben darin, Beobachtungen zu sammeln, zu dokumentieren und an die Projektleitung weiterzugeben.

Die zentrale Frage für den Autor in der Rolle des Evaluators lautete: Welche Informationen können durch die Tutoren erzielt werden? Die folgende stichwortartige Auflistung<sup>298</sup> interessierender Aspekte zeigt, wie vielfältig die Informationen sind, die durch die Tutoren in Erfahrung gebracht werden sollten:

- Verhalten der Studierenden im Tutorium (Beteiligung, Vorbereitung, Zusammenarbeit, Diskussionsbereitschaft): Die Tutoren beobachteten Schwierigkeiten der Studierenden mit den in der Vorlesung vermittelten statistischen Methoden. Die Dozenten griffen diese in den Vorlesungen wiederholt auf, um die Schwächen zu beheben. Weiterhin übermittelten die Tutoren Anfragen der Studierenden bezüglich des Teamworks an die Dozenten.
- Motivation der Studierenden, mit der Lernsoftware zu arbeiten: Die Tutoren beobachteten, dass die Studierenden selbstständig mit den Theoriekomponenten der Lernsoftware arbeiteten, die übrigen allerdings größtenteils nicht nutzten. Die Dozenten reagierten, indem sie z.B. Animationen oder Beispielkomponenten stärker in die Vorlesungen integrierten.
- Probleme der Studierenden im technischen Umgang mit der Lernsoftware: Diese wurden von den Tutoren insbesondere in Bezug auf den Umgang mit dem Statistiklabor identifiziert. Die Dozenten erhöhten den Betreuungsaufwand, gingen in den Vorlesungen verstärkt auf das Statistiklabor und seine Funktionalitäten ein. Auch die Betreuung in der Rechnerstunde wurde intensiviert. Die Dozenten erschienen selbst, um konkrete Probleme zu beobachten und zu beheben. Es stellte sich heraus, dass insbesondere der Umgang mit dem R-Kalkulator Schwierigkeiten hervorrief. Als Reaktion wurde den Tutoren Zeit in den Übungen eingeräumt, die gezielt auf die Vermittlung der Funktionalität des R-Kalkulators verwendet wurde.
- Wo liegt verstärktes Interesse/Lernmotivation - Theorie oder Multimedia: Die Tutoren beobachteten, dass die Studierenden beide Komponenten als gleich wichtig bewerteten. Die Idee der Motivationsvorlesung wurde verstanden, die Notwendigkeit der Theorievorlesung erkannt. Besonderes Interesse wurde der Theoriekomponente von *Statistik interaktiv kompakt* beigemessen. Diese setzten die Studierenden zur Nachbereitung beider Vorlesungen ein. Das Experimentieren im Statistiklabor verstärkte sich im Verlauf des Semesters. Je sicherer die Studierenden im Umgang mit den Funktionalitäten des Labors wurden, desto mehr wuchs ihr Interesse, statistische Probleme eigenständig mit dessen Hilfe zu lösen.
- Bereitschaft der Studierenden, sich auf multimediale Lehre einzulassen, auch wenn Probleme entstehen: Die Tutoren berichteten, dass die Studierenden durch Instabilitäten der Lernsoftware, insbesondere des Statistiklabors, verunsichert wurden. Die technischen Probleme führten dazu, dass die Studierenden das Arbeiten im Labor als schwierig empfanden und versuchten, auf andere Tools wie z.B. MS Excel auszuweichen.
- Wie weit nutzt der Studierende die Software; zeigt er z.B. Interesse an weitergehenden Schritten wie eigenen statistischen Auswertungen unter Verwendung von R? Die Beobachtungen der Tutoren zeigten, dass der R-Kalkulator nur unter starker Betreuung verwendet wurde, das Prinzip des *self-paced-learning*s hier nicht funktionierte. Wurden R-Befehle schrittweise erklärt, waren die Studierenden bereit, sich damit auseinanderzusetzen. Sie wenden vermittelte Befehle und Funktionen auf statistische Probleme an. Der Umgang mit Bibliotheken stellte eine Schwierigkeit dar. In Reaktion auf diese Beobachtungen wurde das Einbinden von Bibliotheken im Statistiklabor vereinfacht.

<sup>298</sup>Diese wird verkürzt dargestellt, da der Leser nicht jede Frage, sondern die Rolle des Aspektes verstehen soll.



Für einen solchen Einsatz sollten die Tutoren bestimmte Qualifikationen mitbringen: Zielstrebigkeit im Studium, vertrauter Umgang mit der Lernsoftware, sehr gute Statistikkenntnisse. Die Evaluationsstudie hat gezeigt, dass diese Fähigkeiten bei weitem nicht ausreichend vorhanden waren.

**Tutorenmanagement:** Die Tutoren wurden vor ihrem Einsatz einer dreitägigen Schulung durch die Projektleitung unterzogen. Sie wurden mit den Zielen des Projektes sowie mit ihrer Rolle und Aufgaben vertraut gemacht. Die Tutoren sollten geeignet auf die Aspekte eingestellt werden, die zur Evaluation der Lernsoftware und des Einsatzkonzeptes interessant waren, um vorstrukturierte Informationen zu sammeln. Den Tutoren wurde in regelmäßigen Abständen ein Beobachtungsbogen ausgehändigt, mit dessen Hilfe sie den Verlauf der Tutorien, das Verhalten der Studierenden gezielt beobachten und dokumentieren sollten.

Es gehörte zu den Aufgaben der Tutoren, die abgegebenen Aufgabenlösungen der Studierenden zu korrigieren. Unter anderem sollte beobachtet werden, ob bestimmte Studierende mehrfach durch schlechte/gute Leistungen auffallen oder ob sich zum Beispiel bei einigen Gruppen ein extremer Leistungsverfall/-steigerung abzeichnete. Außerdem sollte auf diese Art beobachtet werden, ob Bereitschaft seitens der Studierenden bestand, sich mit der Statistik bzw. dem Lösen statistischer Aufgaben selbstständig zu beschäftigen.

Auffällig war, dass die Studierenden trotz des Angebotes der Lernsoftware viele Aufgaben per Hand lösten oder aber die Benutzung anderer Programme, größtenteils MS Excel präferierten. An das Labor wurden die Studierenden langsam durch die Anfertigung der statistischen Reports herangeführt. In der 2. Durchführung der Evaluationsstudie wurden die Studierenden von Anfang an stärker in den Umgang mit dem Statistiklabor eingeführt.

**Bewertung durch die Tutoren** Zur Korrektur der Aufgaben durch die Tutoren ist anzumerken, dass anfangs Kommunikationsschwierigkeiten zwischen der Projektleitung und den Tutoren auftraten. Die Tutoren bewerteten die abgegebenen Aufgaben zwar, setzten zur Unzufriedenheit der Studierenden allerdings lediglich ein r (richtig) oder f (falsch) hinter die abgegebenen Aufgaben. Beschwerden der Studierenden führten dazu, dass nun die Stellen der Aufgabenlösungen, die fehlerhaft waren, durch die Tutoren gekennzeichnet wurden. Dies hatte den Vorteil, die fehlerhafte Stelle zwar zu kennen, jedoch selbst darüber nachdenken müssen, worin die Fehler konkret besteht. Das Präsentieren einer Musterlösung oder eines richtigen Lösungsansatzes seitens der Tutoren hätte den Nachteil gehabt, dass sowohl der Gedankengang wie auch die Begründung zur Lösung fehlten. Genau wie die Beobachtung der Studierenden durch die Tutoren empfiehlt es sich, deren Bewertung der Übungsaufgaben zu steuern. Aus diesem Grunde wurden den Tutoren für jedes Aufgabenblatt Bewertungskriterien angegeben, zum Beispiel in Form eines Bewertungsbogens, der sich auf die Besonderheiten der einzelnen Aufgaben bezog. Auf diese Weise konnte überprüft werden, welche Aufgabentypen den Studierenden besondere Schwierigkeiten bereiteten oder welche sie besonders zum Lernen motivierten. Des Weiteren bekamen die Dozenten einen Überblick darüber, ob der Einsatz der Lernsoftware die Studierenden überforderte oder zu Ablenkungen von dem zu vermittelnden Stoff führte. Ein Bewertungsbogen hat den Vorteil, dass die Bewertung der Tutoren einheitlich, d.h. nach denselben Kriterien vorgenommen wird, es somit zu einer Gleichbehandlung der Studierenden und zu vergleichbaren Ergebnissen parallel betreuter Gruppen kommt.

Am Semesterende wurde eine Gesamtbewertung jedes Projektteilnehmers durch die Tutoren vorgenommen, die in der sogenannten Tutorennote zusammengefasst wurde. Diese setzte sich aus den folgenden nach Relevanz geordneten Aspekten zusammen:

- Qualität der abgegebenen Lösungen der Übungsaufgaben
- Aktivität im Tutorium (Beteiligung an Diskussionen, Demonstration eigener Lösungen, gemeinsames Erarbeiten von Lösungswegen)
- Anwesenheit im Tutorium

**Subjektivität der Bewertung** Die Tutoren vergaben die Noten 1 (sehr gut) - 5 (mangelhaft) sowie einen kurzen Gesamteindruck zu jedem Studierenden. Problematisch muss die Güte der Notenvergabe der Tutoren gesehen werden, d.h. die Tatsache wie gut die Note des Tutors den

tatsächlichen Leistungsstand des Studierenden widerspiegelt.

Vergleicht man, inwiefern die Gesamtnoten der Studierenden übereinstimmen mit den Tutorennoten (Übereinstimmung hoch: 17 von 25), so ist dies nicht präzise, da die Tutorennote zu 35 Prozent in der Gesamtnote enthalten ist. Vergleicht man, um dieses Problem zu umgehen, die Tutorennote mit der Klausurnote der Studierenden, so ist dies inhaltlich nicht sinnvoll. Die Noten sollen gerade unterschiedliche Fähigkeiten der Studierenden überprüfen. Ein Vergleich könnte nur zu Aussagen der Art führen: Gute Studierende sind überall gut, schlechte überall schlecht.

Als Konsequenz aus der fehlenden Aussagekraft der Tutorennoten wurden diese in der 2. Durchführung abgeschafft.

**Aufgetretene Schwierigkeiten** Die Zusammenarbeit mit den Tutoren stellte sich als problematisch heraus. Die Tutoren befanden sich beide im ersten Semester ihres Hauptstudiums, hatten somit selbst viel mit ihrem eigenen Studium zu tun. Das gewünschte Engagement blieb aus, sie taten nichts, was über ihre Pflicht hinausging. Folgende Aussagen von Teilnehmern spiegeln dies wider:

„Hatte man die Aufgaben richtig, war alles o.k. - nur leider war unser Tutor nicht wirklich in der Lage, Aufgaben zu erklären.“

„Das Tutorium war überflüssig, da Grundlagen in der Theorievorlesung gegeben wurden und die reine Aufgabenbesprechung wenig produktiv war.“

„Die gröbere Betrachtung der Probleme der Statistik fand im Tutorium nie statt - oft hätte ich mir eine Ergänzung zu der Theorievorlesung gewünscht“

Als Konsequenz aus diesen Erfahrungen wurde in der 2. Durchführung ein Diplomand eingesetzt, um die Übung zu leiten und die Studierenden mitzubetreuen.

Ein weiteres Problem bestand darin, dass die Tutoren verschiedene Schwerpunkte setzten. Der eine war sehr darauf bedacht, die Studierenden mit R-Befehlen vertraut zu machen. Er war selbst ein leistungsstarker Student und hatte große Schwierigkeiten damit, Verständnisprobleme wie auch mangelnden Arbeitseinsatz der Studierenden nachzuvollziehen und angemessen zu reagieren. Der andere Tutor erklärte detailliert, konnte sich besser in die Situation der Studierenden versetzen, vernachlässigte allerdings die technische Seite.

Es stellte sich heraus, dass Tutoren klare Fragestellungen benötigten, um Beobachtungen zu dokumentieren. Wurde ihnen lediglich ein Bericht über den Verlauf des Tutoriums sowie der Sprechstunden abverlangt, enthielten ihre Ausführungen wenige Informationen.

Die dargestellten Probleme bezüglich der Arbeit mit Tutoren sowie die diesbezüglichen Äußerungen der Studierenden haben gezeigt, wie notwendig eine intensive Personalauswahl und Führung bei der Durchführung einer multimedialen Statistikveranstaltung ist. Die Auswahl sollte gründlich nach fachlichen, aber auch besonders nach persönlichen Kriterien erfolgen, da fehlendes Engagement in einem solchen Projekt deutlichen Schaden anrichten kann.

Im Rahmen der Evaluationsstudie wurde die Konsequenz gezogen, keine Tutoren in dem obigen Sinne einzusetzen. Das Einsatzkonzept sah deshalb den Einsatz eines Diplomanden vor. Wird die Lernsoftware flächendeckend eingesetzt, muss über eine geeignete Vorbereitung und Betreuung von Tutoren nachgedacht werden. Der personelle Mehraufwand darf nicht unterschätzt werden. Die Tutoren benötigen eine stärkere Betreuung als die einer traditionellen Veranstaltung, da sie zahlreiche Aufgaben und Rollen übernehmen. Ideal wäre eine längere Vorbereitungsphase vor Semesterbeginn, in der die Tutoren mit den Zielen der multimedialen Veranstaltung und bereits gewonnenen Evaluationsergebnissen vertraut gemacht werden.

#### 12.2.4 Neue Prüfungsformen

Es wird eine Prüfungsform benötigt, die nicht nur überprüft, ob die Studierenden bestimmte statistische Techniken bzw. die Anwendung statistischer Methoden beherrschen, sondern insbesondere,

ob die Vermittlung statistischen Problemlösens während des Semesters erfolgreich war.<sup>299</sup> In Bielefeld wurde der enge Rahmen der Prüfungsordnung<sup>300</sup> durch einen Experimentierparagrafen für das Fach Statistik erweitert.

§ 29 Experimentierklausel für multimediale Prüfungen:

(1) Bis einschließlich Wintersemester 2003/2004 können im Prüfungsfach Methoden 2 anstelle einer Pflichtveranstaltung in Statistik I und Statistik II je mindestens zwei Wahlpflichtveranstaltungen angeboten werden, von denen jeweils eine der Erprobung neuer Lehr- und Prüfungsformen dienen soll. Diese probeweise angebotenen Veranstaltungen stützen sich auf multimediale Vermittlungsformen der dazugehörigen Stoffgebiete (multimediale Veranstaltungen).

(2) Die Abschlussprüfung zu multimedialen Veranstaltungen erfolgt analog § 5, wobei die Bearbeitung entsprechend aufbereiteter Aufgaben am Rechner mit anschließendem Prüfungsgespräch als weitere Prüfungsform hinzukommt. § 13 Abs. 1 gilt entsprechend.

(3) Jede oder jeder Studierende erhält die Möglichkeit, an einer der angebotenen Wahlpflichtveranstaltungen teilzunehmen. Übersteigt die Zahl der Prüfungsanmeldungen zu multimedialen Veranstaltungen die Zahl der zur Verfügung stehenden Rechnerplätze, ermittelt die Fakultät die Teilnahmeberechtigten per Los. Wer keinen Losplatz erhält, legt die Prüfung in der herkömmlichen Form ab.

Das Einsatzkonzept Ia basierte auf diesem Paragraphen und sah eine Prüfungsform vor, die aus verschiedenen im Folgenden vorgestellten und erläuterten Komponenten besteht. Die Gesamtnote der Projektteilnehmer setzte sich letztendlich wie folgt zusammen:

- zu 45 Prozent aus der Durchschnittsnote der drei statistischen Reports,
- zu 35 Prozent aus der durch die Tutoren vergebene Note
- zu 20 Prozent aus der Multiplechoiceklausur-Note

Betrachten wir die Komponenten der Prüfungsform:

1. Die Projektteilnehmer gaben jede Woche eine **Lösung der Übungsaufgaben** an die Tutoren ab. Diese Einzelleistungen wurden von den Tutoren auf Basis eines von der Projektleitung festgelegten Bewertungsbogens korrigiert. Der Tutor dokumentierte die Leistung jedes Teilnehmers individuell. Die sich aus den abgegebenen Lösungen des Einzelnen ergebende Beurteilung bildete den ersten Teil der Prüfungsleistung.
2. Während des Semesters wurden den Teilnehmern zu festgelegten Terminen drei Aufgaben gestellt. Diese sollten in Form eines **statistischen Reports** in 2er Teams gelöst und dokumentiert werden. Die abgegebenen statistischen Reports wurden anschließend mit den Teilnehmern besprochen. Auf diese Weise sollte überprüft werden, ob und wie die Studierenden aus ihren Fehlern lernten und ob eine Entwicklung hinsichtlich des Verfassens strukturierter Texte in Verbindung mit Ergebnissen des Statistiklabors stattfand, d.h. ob die Hinführung der Studierenden zu statistischem Problemlösen erfolgreich war. Des Weiteren ließ sich durch Befragungen der Studierenden überprüfen, ob die Studierenden zur Teamarbeit in der Lage waren.
3. Der dritte Teil der Prüfung bestand aus einer **Multiplechoiceklausur**. So wurde der Versuch unternommen zu überprüfen, ob der Studierende in der Lage war, statistische Probleme selbständig zu erkennen und zu lösen, ob er statistisches Verständnis besaß, ob er sein Wissen auf reale Probleme übertragen konnte. Sie wurde einer herkömmlichen Klausur vorgezogen, da sie mit weniger Aufwand zu stellen und zu bewerten ist. Die über ihre Möglichkeiten hinausgehenden Kompetenzen wurden mit den statistischen Reports überprüft.

<sup>299</sup>Entwickelt man eine neue Prüfungsform und möchte diese in den Hochschulbetrieb integrieren, so muss gewährleistet sein, dass diese vom zuständigen Prüfungsamt akzeptiert wird. In diesem Zusammenhang entstehen viele organisatorische Probleme.

<sup>300</sup>Vgl. Diplomprüfungsordnung der Fakultät für Wirtschaftswissenschaften, Universität Bielefeld:  
<http://www.wiwi.uni-bielefeld.de/pamt/dpo-so-2002.shtml>

Diese Prüfungsform hatte zum Ziel, eine umfassende Gesamteinschätzung jedes Projektteilnehmers vorzunehmen. Ein wesentliches Ziel bestand darin festzustellen, ob der Einsatz der Multimedia-Software in der statistischen Grundausbildung dazu geführt hat, dass selbstständiges statistisches Arbeiten und Problemlösen beim Studierenden gefördert wurde und er eine positive Haltung gegenüber der Statistik einnahm. Mit der neuen Prüfungsform sollte der Studierende auf verschiedenen Ebenen gefordert und eingeschätzt werden. Gleichzeitig wurde das Instrument, die Lernsoftware, hinsichtlich ihrer Eignung überprüft. Schwächen, wie z.B. fehlende Funktionalitäten des Labors oder fehlende Komponenten (Musterreports, Animationen zum statistischen Report) wurden identifiziert und in den Entwicklungsprozess der Lernsoftware aufgenommen.

Die Multiplechoiceklausur ist eine Komponente der entwickelten Prüfungsform. Sie wird im Folgenden betrachtet, um dem Leser die damit verbundenen Ziele und Schwierigkeiten näherzubringen.

### **Multiplechoiceklausur als Prüfungskomponente**

Ein Blick in die Literatur zeigt, dass es verschiedene Typen von Multiplechoiceklausuren gibt: welche, die bloße Rechentechniken abfragen und als mögliche Ergebnisse unterschiedliche Zahlen angeben, oder welche, die den Aspekt des Rechnens ausblenden und den Schwerpunkt beim statistischen Verständnis setzen. Beide Typen treten auch in Kombination auf. Es erfolgt eine Darstellung der Ideen, die mit der Entwicklung einer speziell für die angestrebten Evaluationsziele entwickelten Multiplechoiceklausur in Verbindung stehen.

Das übergeordnete Ziel der Multiplechoiceklausur bestand darin, zu überprüfen, ob statistisches Verständnis im Laufe des Semesters vermittelt werden konnte. Dieses Ziel wurde in jeder einzelnen Aufgabe, den Lösungsalternativen, aber auch im Mix der Klausuraufgaben berücksichtigt. Statistisches Verständnis wurde im Rahmen der Evaluationsstudie wie folgt definiert:

Statistisches Verständnis ist das Verständnis für die geeignete Anwendung statistischer Methoden zur Lösung realer Probleme, das Verständnis der Zusammenhänge der verschiedenen statistischen Methoden, das Verständnis der Zusammenhänge zwischen der Statistik und allgemeinen Problemstellungen.

Wird eine Multiplechoiceklausur zum Abschluss einer multimedialen Veranstaltung gestellt, so kommt die Frage auf, ob sich der multimediale Aspekt auch in der Klausur niederschlägt bzw. niederschlagen sollte. Multimedia wurde hier eingesetzt, um bestimmte Ziele zu verfolgen. Es wurde als Mittel zum Zweck verstanden und spiegelte sich folglich nicht direkt in der Klausur wider.

**Verschiedene Aufgabentypen und Ziele** Bei dem Entwurf der Multiplechoice-Klausur stellte sich die Frage, ob alle in der Veranstaltung vermittelten Methoden der Statistik durch die Klausur überprüft werden sollten. Anstatt möglichst viele statistische Methoden mit der Klausur abzudecken, wurde darauf geachtet, dass die Klausur eine Reihe von Aufgabentypen beinhaltet, die sich eignen, unterschiedliche Fähigkeiten der Studierenden zu überprüfen. Es wurde zwischen den folgenden Aufgabentypen differenziert:<sup>301</sup>

- Typ A: Aufgaben, die den Umgang mit erlernten statistischen Instrumenten bzw. Methoden trainieren und passende/geeignete statistische Methode zur Lösung eines Problems heranziehen
- Typ B: Aufgaben, die auf das Verständnis der Studierenden abzielen bezüglich der sinnvollen Anwendung von Methoden zur statistischen Problemlösung
- Typ C: Aufgaben mathematischer Art: Herleitungen, Beweise, die zum Ziel haben, die Studierenden behutsam an die mathematische Seite der Statistik heranzuführen, Überprüfung des inhaltlichen Verständnisses statistischer Formeln
- Typ D: Aufgaben, die das Verfassen statistischer Reports trainieren
- Typ E: Aufgaben, die den Umgang mit dem Statistiklabor trainieren
- Typ F: Aufgaben, die das manuelle statistische Problemlösen trainieren

<sup>301</sup>Die Aufgabentypen beziehen sich auf die Übungs- und Klausuraufgaben sowie Reportfragestellungen.

Die verschiedenen Aufgabentypen sollten gewährleisten, dass die Studierenden sowohl die Vorteile des Rechners kennenlernten, als auch in der Lage waren, statistische Berechnungen manuell durchzuführen. Im Vordergrund stand allerdings, ihnen ein Gefühl dafür zu vermitteln, statistische Instrumente zielgerichtet anzuwenden bzw. einzusetzen.

Die Multiplechoiceklausur fand ohne den Einsatz des Rechners statt, schloss somit die Aufgabentypen D und E aus. Im Anhang H stehen Aufgaben aus der Multiplechoiceklausur bereit. Die einzelnen Aufgaben sind hinsichtlich ihres Types gekennzeichnet. Die Schwierigkeit bestand in der Gestaltung der Aufgaben. Als Beispiel kann die Aufgabenstellung des 2. Reports herangezogen werden. Diese betraf die Aufgabentypen A, D und E. Die Schwierigkeit der Studierenden bestand nicht darin, Methoden geeignet anzuwenden, einen statistischen Report zu verfassen oder im Umgang mit dem Labor. Sie konnten aus dem umfangreichen Datenmaterial nicht sinnvoll selektieren - eine Schwierigkeit, die bei der Gestaltung der Aufgabe übersehen wurde.

**Ergebnisse der Multiplechoiceklausur** In der gestellten Klausur überwog der Aufgabentyp A. Der Autor vertritt die Meinung, dass dieser am ehesten dazu geeignet ist, statistisches Verständnis zu überprüfen. Größtenteils zufriedenstellende Ergebnisse lieferten die Aufgaben des Types A. Der Aufgabentyp C wurde größtenteils korrekt gelöst, während Schwierigkeiten bei dem Aufgabentyp B entstanden. Zusammenfassend kann gesagt werden, dass das Gesamtergebnis zwar zufriedenstellend war, man sich seitens der Dozenten jedoch bessere Ergebnisse erhofft hatte. Die durchaus positive Entwicklung der Leistungen der Studierenden in den statistischen Reports sowie das in der Vorlesung gezeigte Interesse, besonders das Engagement, mit dem die Übungsaufgaben gelöst wurden, hatten bessere Ergebnisse versprochen.

### 12.3 Fazit und Grenzen der qualitativen Evaluation

Zum Abschluss der 1. Durchführung wurde den Projektteilnehmern ein Evaluationsbogen ausgehändigt, den alle beantworteten.<sup>302</sup> Das Ziel des Evaluationsbogens bestand darin, die Meinungen der Projektteilnehmer zu erfahren, um daraus Konsequenzen für die 2. Durchführung abzuleiten. Es folgt an dieser Stelle eine Zusammenstellung ausgewählter Zitate bezüglich gekennzeichnete Bereiche:

- Kommunikation: „Man hat gemerkt, dass sich um einen gekümmert wurde.“
- Praxisbezug: „Im Gegensatz zu allen anderen Fächern hat man hier auch mal Probleme aus der Praxis erlebt und nicht nur stupide trockene Theorie gepaukt.“
- Kleine Gruppe: „Aktivität, selbstständiges Denken und Verstehen wurde viel stärker gefordert und gefördert als in den anderen Grundstudiumsveranstaltungen. Dazu hat der Einsatz von Multimedia beigetragen, weil z.B. Reports in diesem Umfang möglich wurden, aber auch die kleine Gruppe.“
- Multimedia: „Meiner Meinung nach hat das Programm *Statistik interaktiv komplett* den Einstieg in die Theorie sehr vereinfacht. Man drückt auf einen Button und schon ist das Histogramm erstellt.“
- Technik: „Da das Programm nur anfangs auf meinem Computer lief (nicht XP tauglich), konnte ich von dem Einsatz von Multimedia nicht sehr profitieren.“
- Belastung: „Teilweise wurde uns zu viel abverlangt, folglich haben wir zeitweise über 20 h in der Woche nur an Statistik gesessen.“
- Tutorium: „Da die Tutorien oft sehr spartanisch waren und die Korrektur der Aufgaben nur mit geringem Engagement durchgeführt wurde, habe ich mich schonmal gefragt, ob die Teilnahme nicht überflüssig ist.“

---

<sup>302</sup>Dieser ist im Anhang F der Arbeit enthalten. Jeder Frage folgt eine Zusammenfassung der Antworten der Studierenden sowie einige interessante Zitate. Es handelt sich um eine rein qualitative Auswertung der Fragen.

- Flächendeckender Einsatz: Die Studierenden merkten an, dass Reformen in Bezug auf die Organisation und Ausstattung der Universitäten für den flächendeckenden Einsatz von Multimedia benötigt werden. „Der Computer in der Vorlesung und im Tutorium war ein Debakel - und das in einer Multimediaveranstaltung!“

Zusammenfassend kann gesagt werden, dass die Projektleitung den Eindruck hatte, dass das statistische Verständnis der Teilnehmer stärker gewachsen ist als in einer konventionellen Statistikveranstaltung. Die Tatsache, dass eine viel stärkere Betreuung als üblich stattfand, schien dabei eine entscheidende Rolle zu spielen. Der gezielte Einsatz der Multimediastsoftware trug einen Teil zu diesem Erfolg bei, ist jedoch nicht überzubewerten. Es war ein großer Vorteil, die Teilnehmer mit Hilfe des Labors in Kombination mit dem Editor an statistisches Problemlösen heranzuführen, obwohl diese Komponenten hinsichtlich ihrer Funktionalität noch nicht ausgereift waren. Die Arbeit mit der Multimediastsoftware machte es möglich, die Teilnehmer mit realen Problemen und Daten arbeiten zu lassen. Sie wurden darauf aufmerksam, dass die Dokumentation und Interpretation erzielter Ergebnisse einen wichtigen Bestandteil des statistischen Problemlösens ausmachen.

Anhand zweier Beispiele wird abschließend erläutert, dass die qualitativen Evaluationsmethoden aufwendig sind und besonders bei größeren Teilnehmerzahlen an Grenzen stoßen.

In der Vorlesung fand keine Beobachtung einzelner Teilnehmer statt, sondern vielmehr eine Gesamtbewertung des Zustandes der Gruppe. Dies stellte sich selbst bei der geringen Anzahl von Teilnehmern als schwierig heraus, da Studierende wenig gewillt sind, Feedback zu geben. In der kleinen Gruppe war es durch die Anwendung der Methode des teilnehmenden Beobachtens möglich, die Gruppe wöchentlich hinsichtlich ihres Kenntnisstandes und statistischen Verständnisses zu bewerten. Die Grenze der qualitativen Evaluation besteht darin, dass auch nur geringfügig größere Gruppen nicht mehr beobachtet werden können. Es bestände die Möglichkeit, bei großen Gruppen nur einige Teilnehmer zu beobachten und die Ergebnisse der Beobachtung zu verallgemeinern. Der Dozent hätte zudem die Möglichkeit, das Kollektiv zu beobachten.

Mit Hilfe des Evaluationsbogens wurde eine abschließende Bewertung der Veranstaltung auf Basis der Äußerungen der Teilnehmer vorgenommen. Die Auswertung war zwar mit einem hohen Arbeitsaufwand verbunden, führte aber zu auswertbaren Evaluationsergebnissen. Es erscheint sinnvoll, eine Kombination qualitativer und quantitativer Fragen anzustreben. Bei Aussagen über das Kollektiv sollte die Evaluation mit einer deutlich größeren Gruppe durchgeführt werden, um verallgemeinerbare Ergebnisse zu erhalten. Um die Evaluation zu standardisieren, wird im folgenden Kapitel ein Evaluationsinstrument entwickelt. Dieses ermöglicht den flächendeckenden Einsatz und Bewertung der Lernsoftware auf Basis des entwickelten Einsatzkonzeptes.

Zusammenfassend konnte nach der 1. Durchführung der Evaluationsstudie festgestellt werden, dass die teilnehmende Beobachtung zu auswertbaren Evaluationsergebnissen führte, allerdings schon bei einer geringen Teilnehmeranzahl mit einem hohen Arbeitsaufwand verbunden war. Als Pilotprojekt war die 1. Durchführung wertvoll, um Erfahrungen und Evaluationsergebnisse zu erzielen.

In der 2. Durchführung der Evaluationsstudie haben Erfahrungen der 1. Durchführung zu konkreten Änderungen im Einsatzkonzept geführt:

- Für die teilnehmende Beobachtung wurde statt Tutoren ein Diplomand eingesetzt.
- Es erfolgte eine stärkere Betreuung der Studierenden bezüglich des Verfassens statistischer Reports (Entwicklung und Implementierung von Musterreports, Nachbesprechungen der Fehler).
- Das konstruktivistische Prinzip des *self-paced-learning*s wurde abgelöst durch das *guided-learning*.
- Die Bereitstellung der Materialien erfolgte strukturierter. Die Organisation wurde durch den Einsatz eines Lernmanagementsystems unterstützt.
- Es wurde ein Evaluationsinstrument entwickelt, das die Evaluation auf Basis eines hohen Stichprobenumfangs ermöglicht. Qualitative und quantitative Evaluationsmethoden können kombiniert werden.

- Das Statistikkolabor wurde stark weiterentwickelt. Dies betraf die Stabilität wie auch eine Funktion zur Erstellung statistischer Reports (Berichterstellungsfunktion).

## 13 Explorative Evaluationsstudie - 2. Durchführung

Dieses Kapitel hat zwei Schwerpunkte. Zunächst werden ausgewählte Komponenten des Einsatzkonzeptes Ib vorgestellt. Anschließend werden die Entwicklung und der Einsatz eines Evaluationsinstrumentes dargestellt. Dieses wurde im Rahmen der 2. Durchführung entwickelt und im Einsatz getestet. Es hat das Ziel, eine flächendeckende Evaluation der Lernsoftware und des Einsatzkonzeptes zu ermöglichen.

### 13.1 Das Einsatzkonzept Ib

#### 13.1.1 Komponenten und Ziele

Das Einsatzkonzept Ib ist die zweite Variante des im Rahmen der Evaluationsstudie entwickelten und evaluierten Einsatzkonzeptes. Die Gestaltung des Einsatzkonzeptes Ib wurde stark beeinflusst durch die Ergebnisse der 1. Durchführung der Evaluationsstudie. Im Folgenden bekommt der Leser Einblicke in ausgewählte Komponenten des Einsatzkonzeptes Ib.

- Verschiedene Veranstaltungen und ein veränderter Grad an Betreuung
- Statistischer Report
- Lernmanagementsysteme
- Weiterentwicklungen von *Statistik interaktiv komplett* und Auswirkungen auf das Einsatzkonzept: das Leitfadensprinzip
- Prüfungsleistung

Das Einsatzkonzept Ib sieht vor, dass die Veranstaltung im Rahmen eines Lernmanagementsystems stattfindet. Im Überblick über ausgewählte Komponenten des Einsatzkonzeptes Ib wird die Idee von Lernmanagementsystemen vorgestellt. Anschließend werden Vor- und Nachteile herausgearbeitet, die sich aus dem Einsatz von *Statistik interaktiv komplett* in dem speziellen Lernmanagementsystem *Blackboard* ergeben haben.

**Verschiedene Veranstaltungen und ein veränderter Grad an Betreuung:** Die Veranstaltung umfasste Vorlesungen, Übungen, betreute Rechnerstunden und Sprechstunden. In allen Veranstaltungen stand die Vermittlung der Statistik mit Hilfe der multimedialen Lernsoftware im Mittelpunkt.

Anders als im Einsatzkonzept Ia wurde zu Beginn der 2. Durchführung eine Aufgabensammlung herausgegeben. Der Unterschied zu wöchentlichen Aufgabenzetteln besteht darin, dass der Studierende einen Überblick darüber hat, was ihn im Semester erwartet. Außerdem ist die Aufgabensammlung nach denselben Themengebieten wie die Vorlesung unterteilt, was zusätzliche Orientierung bietet. Die Aufgabensammlung stellt einen Pool dar - es steht dem Dozenten frei, welche Aufgaben er direkt in der Vorlesung mit den Studierenden löst und welche er lediglich zu Übungszwecken verwendet. Auch der Studierende hat die Möglichkeit, selbstständig Aufgaben zu jedem Zeitpunkt zu üben. Dies gibt dem Studierenden größere Freiheiten und Möglichkeiten, selbstständig anhand von Übungsaufgaben Wissen aus der Vorlesung zu vertiefen. Er kann aktiv werden, ohne dies unter dem Druck der Lösungsabgabe zu tun, und auswählen, welche Inhalte er zusätzlich zu den obligatorischen Aufgaben üben möchte. Auch dies ist wieder auf den konstruktivistischen Aspekt des freien Lernens und der selbstständigen Wissenskonstruktion zurückzuführen.

**Grad der Betreuung:** Die Erfahrungen der 1. Durchführung der Evaluationsstudie zeigten, dass das Prinzip *self-paced-learning* nicht funktioniert. Im Einsatzkonzept Ib wurde daher das Prinzip *guided-learning* verfolgt. Der Studierende wurde vom Dozenten und Tutor durch die Statistik und die Veranstaltung geleitet. Dies wirkte sich auch auf die Gestaltung der Materialien (Komponenten

der Lernsoftware, Aufgabensammlung, Drehbuch, Leitfaden) aus. Folgt man dem Konstruktivismus, so ist dies mit einer expliziten Kontrolle der Studierenden schwer vereinbar. Die Erfahrungen zeigten, dass diese jedoch unumgänglich ist. Die Abgabe von Übungsaufgaben war deshalb obligatorisch. Die Teilnahme an den Vorlesungen, Übungen und betreuten Rechnerstunden war nicht obligatorisch. Im Folgenden wird der Übergang vom *self-paced-* zum *guided-learning* am Beispiel statistischer Reports vermittelt:

**Statistischer Report:** Die positiven Erfahrungen mit dem statistischen Report als Instrument zur Vermittlung statistischen Verständnisses und Problemlösens haben dazu geführt, einen Musterreport als multimediale Komponente in die Lernsoftware zu integrieren. Im Einsatzkonzept Ia erhielten die Studierenden keine Musterreports, erstellten eigene Reports nach dem Prinzip des *self-paced-learnings*. Die Erfahrungen zeigten, dass es den Studierenden schwer fiel, den Zugang zum eigenständigen Verfassen statistischer Reports zu finden. Wiederholt wurde die Bitte an die Projektleitung gestellt, ein Beispiel zur Verfügung zu stellen, das bei der Erstellung als Muster zur Verfügung steht. Um das schrittweise Vorgehen beim Erstellen eines Reports zu vermitteln, wurde ein Report in der neusten Version von *Statistik interaktiv komplett* verankert und schrittweise erstellt. Zusätzlich wurde eine sogenannte Coachingkomponente entwickelt und implementiert, die allgemeine Aspekte der Reporterstellung vermittelt. Die Entwicklung der Coachingkomponente folgte dem Prinzip des *guided-learnings*. In einer Animation wurde die schrittweise Erstellung des Musterreports außerdem visualisiert.

**Evaluationsergebnisse bezüglich der statistischen Reports:** Nach Vermittlung der univariaten Datenanalyse wurde von den Studierenden der erste Report erstellt, der Teil der Prüfungsleistung war. Den Studierenden wurden drei Datensätze zur Verfügung gestellt. Sie sollten die Daten mit den erlernten statistischen Methoden analysieren und einen statistischen Report erstellen. Die Aufgabenstellung wurde wie auch in der 1. Durchführung bewusst allgemein gehalten, um den Studierenden keine Entscheidung bezüglich der Methodenwahl abzunehmen. Die Reports wurden in Zweierteams angefertigt. Ein Großteil der Reports war übersichtlich strukturiert. In einer Problemstellung grenzten die Teams das Problem ein, wählten im nächsten Schritt begründet statistische Methoden aus. Es folgte die Analyse der Daten im Labor. Die Ergebnisse wurden in Reportform dokumentiert. Dieser wurde mit der Berichterstellungsfunktion erstellt. Es erfolgte ein Rückschluss zur Realität.

Diese Struktur wurde weitestgehend eingehalten. Schwächen bestanden bei einigen Reports bei der Problembeschreibung sowie dem Rückschluss auf die Realität, d.h. es wurde viel analysiert, allerdings nicht dokumentiert, beschrieben, begründet oder interpretiert. Das Spektrum der angewendeten Methoden reichte von der Erstellung von Häufigkeitstabellen und vergleichender Histogramme und Boxplots über die Berechnung von Lage- und Streuungsmaßzahlen bis hin zu Tortendiagrammen. Der Zusammenhang zwischen Histogramm und Boxplot wurde größtenteils erkannt und erläutert. Viele Reports zeichneten sich dadurch aus, dass die Studierenden die Daten genau und kritisch betrachteten, bevor sie mit der Analyse begannen. Das in der Vorlesung demonstrierte *Look at your data!* wurde angewendet. Alle Teams haben die Datensätze zunächst betrachtet und dann begründet Bereiche (Zeilen, Spalten) zur näheren Analyse ausgewählt. Auch positiv hervorzuheben ist die Einbringung von eigenen Ideen bei der Interpretation und die Vermutungen darüber, auf welche inhaltlichen Aspekte einzelne Ergebnisse der Datenanalyse zurückzuführen sein könnten. Bei dem zweiten statistischen Report wurden analoge Beobachtungen erzielt.

Die Bearbeitung des dritten Reports erfolgte nicht wie bei den beiden ersten Reports im Team. Das Ziel bestand darin, eine abschließende Einzelleistung jedes Teilnehmers zu erhalten. Die Situation<sup>303</sup> wurde so gewählt, dass viele im Semester erlernte statistische Methoden auf die Daten angewendet werden konnten. Dies wurde von einem Großteil der Teilnehmer genutzt. Sie erstellten Häufigkeitstabellen, argumentierten über Klassenbreiten, hoben den stetigen Charakter der Daten hervor und erstellten vergleichende Histogramme und Boxplots, Dichte-, Verteilungsfunktionen und

---

<sup>303</sup> Analyse zweier Datensätze, die die Jahresdurchschnitte der Niederschläge für NYC für den Zeitraum 1869 bis 1977 enthalten und sich voneinander unterscheiden



Streudiagramme.

Zusammenfassend kann festgestellt werden, dass die schwachen Teilnehmer im Verlauf des Semesters schwach geblieben sind, allerdings eine leichte Verbesserung hinsichtlich ihrer Dokumentation stattgefunden hat. Ansonsten kann von einer deutlichen Verbesserung aller anderen Gruppen gesprochen werden, was sowohl die begründete Auswahl und Anwendung statistischer Methoden als auch die Dokumentation des statistischen Arbeitens und die erzielten Ergebnisse betraf.

Vergleicht man die Leistungen und Entwicklungsprozesse der Studierenden mit denen der 1. Durchführung, so ist eine deutliche Verbesserung erkennbar. Die teilnehmende Beobachtung hat zu dem Ergebnis geführt, dass die Studierenden die Reporterstellung und besonders Strukturierung durch die neu entwickelten Komponenten (Coaching, Animation, Musterreports) besser erlernten. Die Studierenden setzten alle Komponenten ein, während sie ihre eigenen Reports erstellten. Bei der Reporterstellung spielte eine weitere Komponente des Einsatzkonzeptes Ib eine entscheidende Rolle: das Lernmanagementsystem Blackboard.

**Lernmanagementsysteme:** Eine weitere Komponente des Einsatzkonzeptes Ib ist das Lernmanagementsystem Blackboard. Die Idee von Lernmanagementsystemen wird an dieser Stelle kurz vorgestellt.

Lernmanagementsysteme oder Lernplattformen<sup>304</sup> sind internetbasierte Lernumgebungen, in denen Studierende und Lehrende ohne spezielle technische Kenntnisse (online) lernen/lehren können. Durch den Einsatz solcher Lernplattformen besteht die Möglichkeit, traditionelle und multimediale Veranstaltungskomponenten zu kombinieren, einen umfassenden Rahmen um die Veranstaltung zu legen. Kursinhalte können in verschiedenen Formen aufbereitet und bereitgestellt werden. Des Weiteren können diese übersichtlich strukturiert und sortiert werden. Eine mögliche Definition lautet:

„Lernmanagementsystem (LMS), englisch: Learning Management System, alternativ wird auch der Begriff „Lernplattform“ verwendet: ein Softwaretool, auf das über das Internet/Intranet zugegriffen werden kann und das bestimmte Funktionalitäten wie die Administration der Lernenden und der Lerninhalte möglich sind. Im allgemeinen werden in der Lernumgebung bestimmte Kommunikationsmöglichkeiten wie z.B. Tutor-Mail, Newsgroups/Diskussionsforum, Dokumentenpool bereitgestellt. Das LMS ist die Schnittstelle zwischen dem Tele-Tutor und den Lernenden (untereinander).“<sup>305</sup>

Die Vorteile gegenüber der individuellen Organisation der Veranstaltung durch den Dozenten ist eine standardisierte Struktur und Vorgehensweise bei der Planung und Durchführung einer Veranstaltung. Im sich anschließenden Kapitel wird die Komponente *Lernmanagementsystem* ausführlich dargestellt und diskutiert, da ihr eine wichtige Rolle in der 2. Durchführung der Evaluationsstudie zukam.

**Statistik interaktiv komplett:** Im Einsatzkonzept Ib wurde eine im Vergleich zur 1. Durchführung weiterentwickelte Version bzw. ein Prototyp<sup>306</sup> der multimedialen Lernsoftware *Statistik interaktiv komplett* evaluiert. Die Entwicklung neuer Komponenten (Musterreport, Coachingkomponente, Animation) sowie der Einsatz in einem Lernmanagementsystem haben parallel zu Änderungen des Einsatzkonzeptes geführt. Die Studierenden wurden durch die neuen Komponenten stärker an das Verfassen statistischer Reports herangeführt. Um den Studierenden gezielt durch die Fülle angebotener Materialien zu leiten, wurde eine weitere Komponente in das Einsatzkonzept Ib integriert.

**Der Leitfaden:** Für jede einzelne Vorlesung existierte ein Leitfaden, anhand dessen die Vorlesung schrittweise durchgeführt wurde. Der Leitfaden bildete das Gerüst der jeweiligen Vorlesung. Auf einem maximal zweiseitigen Dokument wurden traditionelle und multimediale Elemente der Vorlesung direkt miteinander verbunden. Ein Leitfaden enthält Verweise auf alle Materialien: mul-

<sup>304</sup>Diese Begriffe werden synonym verwendet.

<sup>305</sup>[www.elearnconcept.de](http://www.elearnconcept.de).

<sup>306</sup>Vgl. 10, Prototyp III.

timediale Lernmodule bzw. einzelne Komponenten, relevante Literaturquellen, Teile des herausgegebenen Skriptes, Übungsaufgaben der zu Semesterbeginn verteilten Aufgabensammlung, speziell für die jeweilige Vorlesung entwickelte Statistiklaborseiten. Der Leitfaden führt den Studierenden durch diese vom Dozenten ausgewählten, teilweise individuell erstellten Materialien, gibt zudem eine vom Dozenten empfohlene Vorgehensweise/Reihenfolge vor. Der Leitfaden wird innerhalb der Vorlesung schrittweise durchgearbeitet und steht den Studierenden im Anschluss an die Vorlesung im Lernmanagementsystem Blackboard zur Verfügung. Ein Vorteil des sogenannten Leitfadenprinzips besteht in der organisierten und nachvollziehbaren Navigation durch die verschiedenen Komponenten der multimedialen Lernmodule. Verlinkungen führen den Studierenden direkt von Stellen des Leitfadens zu den jeweiligen Komponenten von *Statistik interaktiv komplett*. Dies wird durch die Organisation der Veranstaltung im Lernmanagementsystem Blackboard ermöglicht, in dem Leitfaden und Lernmodule bereitgestellt werden. Dadurch ist es möglich, in Form aktiver Verlinkungen vom Leitfaden auf die Komponenten der Lernmodule zu verweisen. Für den Studierenden ergeben sich aus dem Leitfadenprinzip folgende Vorteile:

- Er kann sich während der Vorlesung allein auf die Inhalte der Vorlesung konzentrieren, muss weder Komponenten der Lernmodule noch deren Reihenfolge dokumentieren.
- Er verliert sich beim Nacharbeiten der Vorlesung nicht in der Fülle der Materialien, sondern navigiert organisiert nach Vorgabe des Leitfadens.
- Er hat die Möglichkeit, im Statistiklabor auf Seiten des Dozenten Modifikationen vorzunehmen und Variationen auszuprobieren, da auch das Statistiklabor aus dem Leitfaden heraus aufgerufen werden kann.

Der Leitfaden wurde als Konsequenz der 1. Durchführung der Evaluationsstudie entworfen. Während dieser wurde deutlich, dass die Studierenden durch zu viele Möglichkeiten der Navigation und zu viele verschiedene Materialien überfordert waren. Die 2. Durchführung hat gezeigt, dass Aktivität und Wissenskonstruktion seitens des Studierenden durch das Leitfadenprinzip hervorgerufen und gefördert wird.

**Prüfungsleistung:** Anders als in der herkömmlichen Grundstudiumsveranstaltung Statistik bestand die Prüfungsleistung nicht in einer Klausur. Die Prüfungsleistung setzte sich aus der Anfertigung dreier statistischer Reports (20% - 20% - 40%) während des Semesters zusammen. Dabei wurden die ersten beiden Reports in Zweiergruppen angefertigt, bei dem dritten Report handelte es sich um eine Einzelleistung jedes Teilnehmers. An diesen war ein ca. 20 minütiges Gespräch/eine mündliche Prüfung gebunden. Zu 20 % wurden die wöchentlich abzugebenen Übungsaufgaben in die Endnote eingerechnet. Der mündlichen Prüfung kam die Rolle zu, die Fähigkeit zur Wissenskonstruktion exemplarisch zu überprüfen. Die Studierenden wurden mit ihren eigenen im Semester angefertigten Reports konfrontiert. Auf diesen aufbauend wurden Fragen zu gewählten Verteilungsmodellen, ihren Parametern und alternativen Methoden und Interpretationen gestellt. Die Studierenden mussten während der Prüfung zeigen, inwiefern sie altes Wissen mit neuen Aspekten verknüpfen konnten.

Durch die Methode des teilnehmenden Beobachtens überprüften die Evaluatoren, inwieweit statistisches Verständnis sowie Wissenskonstruktion geschaffen werden konnten. Bei jedem Team wurde der konstruktivistische Prozess im Rahmen der mündlichen Prüfung im Kleinen angestoßen. Der Prüfer durfte aus diesem Grunde nicht zu viele Hinweise geben, sollte den Studierenden die Möglichkeit geben, über alternative Vorgehensweisen laut nachzudenken, alternative Lösungsstrategien zu erarbeiten.

### 13.1.2 Integration eines Lernmanagementsystems

Eine zentrale Komponente des Einsatzkonzeptes Ib war das Lernmanagementsystem Blackboard. Die Lernmodule wurden im Rahmen von Blackboard eingesetzt und evaluiert. Die Idee von Lernmanagementsystemen wurde bereits dargestellt. In diesem Abschnitt findet eine Diskussion in drei Schritten statt:

1. Darstellung der Informationsverarbeitung in der Veranstaltung: Es werden die Informationsströme zwischen den an der Veranstaltung beteiligten Personen dargestellt.
2. Ableitung der sich aus den Informationsströmen ergebenden Anforderungen an ein Lernmanagementsystem: Zunächst wird von einem konkreten Lernmanagementsystem abstrahiert, um in einem ersten Schritt generelle Anforderungen an ein Lernmanagementsystem herauszuarbeiten.
3. Erfahrungen mit dem Lernmanagementsystems Blackboard: Es wird berichtet, zu welchen Erfahrungen die Nutzung von Blackboard geführt hat, inwiefern es den Anforderungen des Einsatzkonzeptes Ib gerecht wurde.<sup>307</sup>

### **Darstellung der Informationsverarbeitung in der Veranstaltung**

Für die folgende Betrachtung werden die Handlungen der an der multimedialen Veranstaltung Beteiligten in Bezug auf Informationen und deren Verarbeitung betrachtet. Diese Perspektive ist für die Spezifikation der Anforderungen an ein Lernmanagementsystem mitentscheidend, weil so ein Teil der zu fordernden Eigenschaften genauer beschrieben werden.

Als handelnde Beteiligten sind der Dozent, der Tutor<sup>308</sup> und der einzelnen Studierende sowie die Studierenden als Lerngruppe zu identifizieren. Im Folgenden werden die Tätigkeiten der einzelnen Beteiligten sowie ihr nötiges Zusammenwirken und Interagieren dieser betrachtet.

Der **Dozent** hat folgende Tätigkeiten im Rahmen der Veranstaltung zu erledigen:

- **Materialsammlung:** Der Dozent erstellt seine Materialien, z.B. Statistiklaborseiten, die Aufgabensammlung und die Leitfäden. Diese lagert er in seinem, nur ihm zugänglichen Bereich.
- **Materialaufbereitung:** Die Materialien werden individuell für die Veranstaltung aufbereitet. Es werden z.B. Statistiklaborseiten für spezielle Vorlesungen vorbereitet. Am Beispiel der Materialien Leitfäden und Lernmodule zeigt sich, dass erst Verknüpfungen/Verlinkungen das Material für die Veranstaltung einsetzbar machen.
- **Materialsortierung:** Das Material wird den einzelnen Veranstaltungsterminen zugeordnet und dementsprechend sortiert. Auf diese Sortierung greift der Dozent während seiner Veranstaltungen zurück.
- **Materialverteilung:** Das Statistiklabor wird an die Studierenden verteilt. Entweder als CD oder per Download. Skripte, Lernmodule, Leitfäden, Aufgabensammlung und Datensätze werden an die Studierenden verteilt und stehen ihnen während der gesamten Veranstaltung zur Verfügung. Die Evaluationsbögen werden zu Semesterbeginn und -ende an die Studierenden verteilt. Daraus ergeben sich verschiedene Formen der Materialverteilung. Das Material wird an verschiedene Empfänger verteilt, und es wird den Empfängern entweder dauerhaft, nur ab bestimmten Zeitpunkten oder für bestimmte Zeiträume zur Verfügung gestellt.
- **Materialempfang:** Studierende geben ihre statistischen Reports im System an den Dozenten oder Tutor ab. Dieser verwaltet er in seinem persönlichen Bereich.

Dem **Tutor** sind folgende Tätigkeiten im Einsatzkonzept Ib zgedacht.

- **Materialempfang:** Der Dozent stellt dem Tutor Materialien zur Veranstaltung, z.B. Statistiklaborseiten, zur Verfügung. Die Studierenden geben bei ihm Bearbeitungen der wöchentlichen Aufgaben ab. Folglich empfängt der Tutor von allen Beteiligten Material.
- **Materialverteilung:** Dem Dozenten liefert der Tutor Auswertungen und Berichte über die Aufgabenbearbeitung der Studierenden. Er übergibt dem einzelnen Studierenden Korrekturen der abgegebenen Ausarbeitungen. Den Teams stellt er Hilfestellungen und Materialien für die Tutorien zu Verfügung. Somit verteilt der Tutor Materialien an den Dozenten, den einzelnen Studierenden und die Teams.

<sup>307</sup> Ausführungen zu technischen Problemen im Einsatz und diesbezüglich entwickelten Lösungsmöglichkeiten wurden in den Anhang I der vorliegenden Arbeit ausgelagert.

<sup>308</sup> Dies ist im Einsatzkonzept ein Diplomand, der im Rahmen des Forschungsprojektes arbeitet.

- **Materialverwaltung:** Der Tutor verwaltet und speichert die Ausarbeitungen der Studierenden und die Materialien des Dozenten in seinem persönlichen Bereich.

Der **Studierende bzw. die Studierenden als Team** haben folgende Tätigkeiten:

- **Materialempfang:** Der Studierende erhält vom Dozenten, die für ihn bereitgestellten Veranstaltungsmaterialien. Der Tutor gibt ihm die Korrekturen seiner Bearbeitungen der Aufgaben. Andere Studierende liefern während der Teamarbeit Teillösungen und aufbereitete Materialien. Damit ist der Studierende in der Lage von allen Beteiligten Material in Empfang zu nehmen.
- **Materialverteilung:** Dem Dozenten übergibt der Studierende bearbeitete Evaluationsbögen und die erstellten statistischen Reports. Der Tutor erhält seine wöchentliche Bearbeitung der Aufgaben. Für die Zusammenarbeit mit anderen Studierenden gibt er Teillösungen und andere Materialien weiter. Der Studierende gibt Materialien aus seinem persönlichen Bereich an Studierende, den Tutor oder den Dozenten weiter.
- **Materialverwaltung:** Der Studierende organisiert die Materialien in seinem persönlichen Bereich.

Neben den dargestellten Tätigkeiten, die sich mit Materialien beschäftigen, tauschen die Beteiligten in den gemeinsamen Terminen noch mündlich Informationen aus. Aus den Tätigkeiten ergeben sich folgende Daten-/Informationsströme zwischen den an der Veranstaltung Beteiligten:

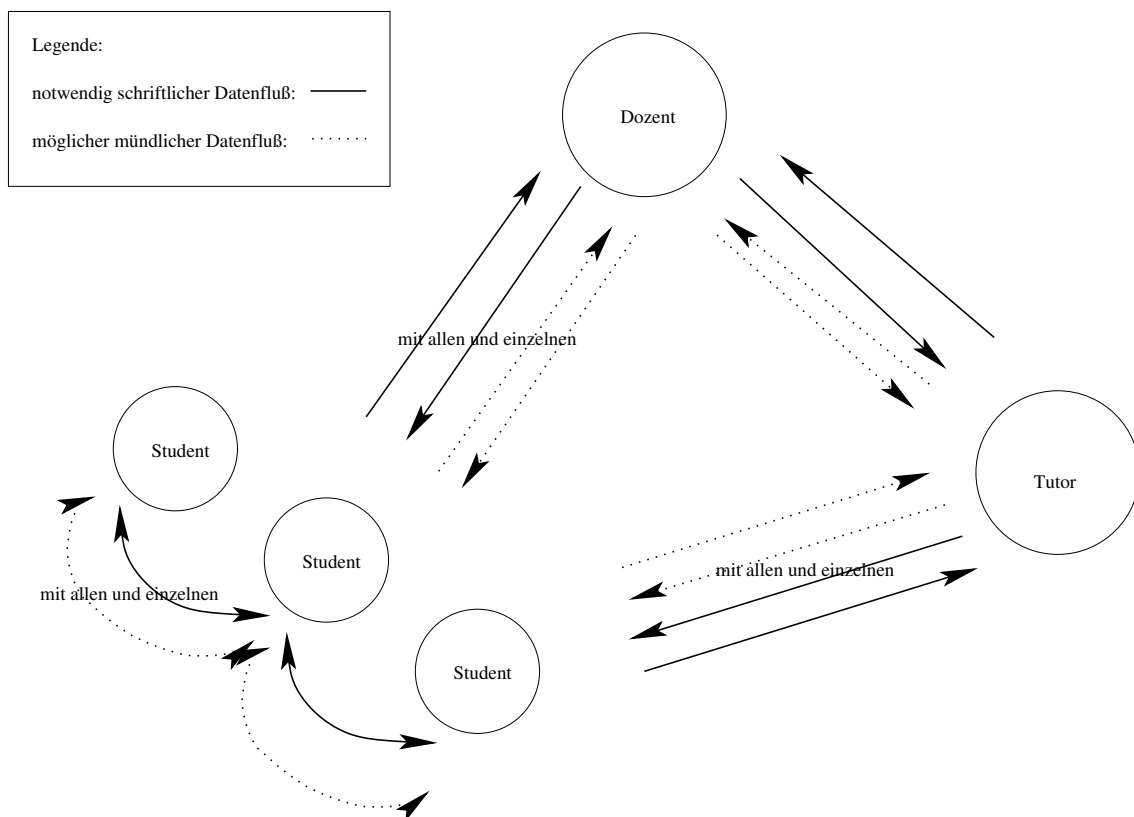


Abbildung 66: Informationsströme zwischen den Beteiligten der Veranstaltung

Die Betrachtung des Einsatzkonzeptes Ib in Hinsicht auf den Informationsfluss hat ergeben, dass zwischen allen Beteiligten der Veranstaltung ein Austausch schriftlicher sowie mündlicher Informationen stattfindet. Dabei benötigen die Beteiligten die Kontrolle über den Zeitpunkt sowie die

Adressierung der Informationen. Jeder hat einen privaten Bereich, in dem er seine Informationen erarbeitet und organisiert aufbewahrt. Neben diesem Bereich kann jeder in einem gemeinsamen, öffentlichen Bereich Daten zur Verfügung stellen und abrufen.

Aus diesen im Einsatzkonzept Ib vorgesehenen Informationsströmen ergeben sich Anforderungen an ein Lernmanagementsystem. Diese werden im Folgenden abgeleitet.

**Anforderungen an ein Lernmanagementsystem:** in sechs Bereichen werden idealtypische Anforderungen an ein Lernmanagementsystem überblickartig dargestellt. Das Lernmanagementsystem soll in der Handhabung und Benutzung einfach sein. Die von dem Lernmanagementsystem übernommenen Aufgaben müssen für die beteiligten Personen wesentlich einfacher werden. Durch das System sollen Kapazitäten freigesetzt werden, die den Lern- und Lehrerfolg der Veranstaltung steigern. Der Einsatz eines Lernmanagementsystems ist mit dem eines Werkzeuges zu vergleichen. Das Werkzeug unterstützt seinen Benutzer bei seinen Tätigkeiten. Der Einsatz eines Werkzeuges sollte die Arbeitsweisen jedoch nicht derart verändern, dass eine Reihe weiterer Hilfsmittel oder /-konstruktionen notwendig werden und der Aufwand erhöht wird.

1. **Rahmenumgebung:** Das Lernmanagementsystem soll für alle Komponenten einen Rahmen bilden. Es soll sicherstellen, dass keine zusätzlichen infrastrukturellen Maßnahmen am Veranstaltungsort ergriffen werden müssen.
2. **Veranstaltungsorganisation:**  
Mit Hilfe des Lernmanagementsystems sollen sich alle organisatorischen Aufgaben der Veranstaltung bewältigen lassen. Das beginnt mit der Ankündigung der Veranstaltung im Vorfeld bis hin zur Anmeldung der Studierenden zur Veranstaltung und zu den Prüfungen. Aktuelle Mitteilungen sollen im Lernmanagementsystem an die Gruppe verteilt werden. Das Lernmanagementsystem soll die Möglichkeit bieten, die verschiedenen Rollen der Teilnehmer zu unterstützen. Die Dozenten, Tutoren und Studierenden sollen mit den Rechten und Fähigkeiten ausgestattet werden, die das Einsatzkonzept Ib für sie vorsieht.
3. **Kommunikation:** Das Lernmanagementsystem stellt eine Kommunikationsumgebung her. Diese fördert die Kommunikation der Studierenden untereinander sowie die Kommunikation zwischen Studierenden und Dozenten. Es gibt verschiedene Arten von Kommunikation: Email, Diskussionsforen, Austausch von Materialien. Letzteres steht im Mittelpunkt. Dies betrifft besonders die Komponente *Statistischer Report*. Der Studierende kann sich im Lernmanagementsystem sowohl die Aufgabenstellung als auch die benötigten Daten abholen und diese in seinem individuellen Bereich bearbeiten. In diesem Bereich kann er seine Zwischenergebnisse speichern und ausgewählte Materialien organisieren. Teilschritte seiner Lösung kann er mit seinem Team innerhalb des Lernmanagementsystems auf Statistiklaborseiten diskutieren und austauschen. Die fertiggestellte Lösung wird zur Korrektur an den Dozenten/Tutor weitergeleitet. Dieser korrigiert den Report im System und übergibt die korrigierte Version wieder an den Studierenden bzw. das Team. Diese Transaktion betrifft das Statistiklabor, die Berichterstellungsfunktion und individuelle Bereiche. Der Studierende soll seine Bereiche im Lernmanagementsystem als Arbeitsumgebung für die Veranstaltung gestalten können. Es muss gewährleistet sein, dass er seine Arbeitsergebnisse und -zwischenritte im Lernmanagementsystem verwalten kann.
4. **Materialien:** Vorlesungsmaterialien, z.B. vom Dozenten angefertigte Vorlesungslaborseiten, sind vom Studierenden innerhalb des Lernmanagementsystems zugänglich. Er muss auf diese zugreifen und sie modifizieren können. Dies betrifft sowohl das Statistiklabor, als auch einen eigenen Bereich des Studierenden sowie weitere Vorlesungsmaterialien, wie z.B. das Skript oder den Leitfaden. Die Modifikationen dürfen nur innerhalb seines Bereiches möglich sein, damit die ursprüngliche Version des Dozenten erhalten bleibt. Der Studierende kann sich innerhalb seines Bereiches ein eigenes Skript erstellen, bestehend aus ausgewählten Materialien des Dozenten und eigenen oder modifizierten.
5. **Zugriff:** Der Studierende hat sowohl von zu Hause als auch von den Rechnern des Hochschulrechenzentrums Zugriff auf das Lernmanagementsystem und somit alle Kursmaterialien.

Auch im Hörsaal besteht diese Zugriffsmöglichkeit. Es existiert ein *Notfallplan*, der mögliche Ausfallszenarien und Konsequenzen berücksichtigt und Lösungen anbietet. Dieser gewährleistet, dass bei technischen Problemen zumindest während der Vorlesung und Übung weiterhin Zugriff auf die multimedialen Lernmodule sowie Vorlesungsmaterialien besteht.<sup>309</sup>

- 6. Verwaltung von Kursmaterialien:** Materialien der Veranstaltung können auf verschiedene Art eingebracht, freigegeben und verwaltet werden. Dem Dozenten soll ermöglicht werden, seine Veranstaltung im zeitlichen Ablauf zu organisieren. Dies ist auch bei der Bereitstellung von Musterlösungen für Reports interessant. Die Musterlösung wird dem Studierenden vom System erst dann bereitgestellt, wenn der Dozent diesen Mechanismus freigibt oder ein Termin abgelaufen ist.

### Lernmanagementsystem Blackboard – Bewertung und Erfahrungen

Das Lernmanagementsystem Blackboard wurde in der 2. Durchführung der Evaluationsstudie eingesetzt und hinsichtlich seiner Eignung überprüft. Der Großteil der im voranstehenden Abschnitt aufgeführten idealtypischen Anforderungen wurde von Blackboard erfüllt. Daher hat der Einsatz von Blackboard viele Verbesserungen bewirkt. Die Kommunikation zwischen den Beteiligten der Veranstaltung wurde deutlich gefördert. Die Organisation der Fülle an Materialien und besonders deren Verknüpfung wurde durch Blackboard ermöglicht. Trotz des Erfolges ist die Integration in das Einsatzkonzept Ib kritisch zu sehen. Sie war mit einem hohen Zeit- und Personalaufwand verbunden. Es entstanden zudem neue Probleme, z.B. die Möglichkeit des technischen Ausfalls des Systems während der Veranstaltung. Auch die Entwicklung eines detaillierten *Notfallplans* war mit viel Aufwand verbunden. Vor dem Einsatz von Blackboard stellten sich die Projektbeteiligten die Frage, was man genau braucht und ob die Anforderungen durch alternative Infrastrukturen gedeckt werden könnten. Einige Projektpartner vertraten die Position, der benötigte Informationsaustausch könne auch über das Internet, z.B. eine projektinterne Webseite erfolgen. Diese ist im Gegensatz zur Nutzung von Blackboard mit keinen zusätzlichen Kosten für die Hochschule verbunden. Die Entscheidung fiel nach der Diskussion der oben dargestellten Anforderungen für den Einsatz des Lernmanagementsystems Blackboard. Insbesondere die Verknüpfung der Leitfäden und Lernmodule sowie die bei der Reporterstellung benötigten und ermöglichten Informations-/Datenströme unterstützten die Entscheidung für Blackboard. Zudem würde die Größe der Lernmodule ein Problem für die Bereitstellung und Nutzung über eine Webseite darstellen.

Ein Defizit des Lernmanagementsystem Blackboard liegt darin, dass es zwar möglich ist, zwischen einzelnen Materialien zu verlinken, dass allerdings nicht die Möglichkeit besteht, in bestimmte Materialien hinein zu verweisen. Im Leitfaden wird der Studierende z. B. in den statistischen Kapiteln auf die entsprechenden multimedialen Lernmodule zu diesen Themen verwiesen. Es wäre wünschenswert, zwischen Leitfaden und den Theoriekomponenten der entsprechenden Lernmodule zu verlinken. Dies ist technisch nicht möglich. Der Studierende wird im Leitfaden auf die Lernmodule verwiesen und kann die Theoriekomponenten gesondert aufrufen.

Eine weitere Anforderung, die technisch von Blackboard nicht erfüllt werden kann, bezieht sich auf den Austausch von Materialien zwischen den Studierenden und dem Dozenten. Es ist zunächst vorteilhaft, dass in Blackboard die Möglichkeit besteht, in Zweiergruppen zu arbeiten. Dies ist besonders für das Verfassen der statistischen Reports von Interesse. Die Gruppenmitglieder besitzen einen gemeinsamen Ordner zum Dateiaustausch. Das Defizit liegt darin, dass sie als Gruppe keine Lösung an den Dozenten schicken können. Dies kann nur aus dem individuellen Bereich eines Gruppenmitgliedes stattfinden.

Da Blackboard viele verschiedene Nutzungsmöglichkeiten bietet, wurde in Bielefeld ein Manuell für Studierende und eines für Dozenten entwickelt. Diese enthalten Beschreibungen der in Bielefeld für die Evaluationsstudie gewählten Gestaltung von Blackboard. Ihr Schwerpunkt liegt auf der Erarbeitung der statistischen Reports im Team. Die diesbezügliche Kommunikation zwischen den beteiligten Studierenden, dem Tutor und dem Dozenten war ausschlaggebend für die Nutzung der von Blackboard bereitgestellten Möglichkeiten. Das Lernmanagementsystem kann auch alternativ

<sup>309</sup>Es wurde ein Notfallplan für die Evaluation des Einsatzkonzeptes Ib entwickelt. Dieser steht dem Leser im Anhang I der vorliegenden Arbeit zur Verfügung.

eingesetzt werden. Denkbar wäre zum Beispiel, es als reines Materialbereitstellungslager zu nutzen. Ist dies der Fall, stellt sich allerdings die Frage, ob ein Laufwerk nicht mit erheblich weniger Aufwand denselben Nutzen erbringen würde.

### 13.1.3 Zielkompetenzen und Instrumente zur Zielerreichung

Der Einsatz multimedialer und anderer Veranstaltungskomponenten verfolgte das Ziel, die Statistikausbildung zu verbessern. Den Studierenden sollten statistische Kompetenzen vermittelt werden, die die bisherige Statistikausbildung nicht oder nur zum Teil vermitteln konnte. Die im folgenden zusammengestellten Fähigkeiten enthalten diese statistischen Kompetenzen, stellen das Ziel unseres Einsatzes neuer Medien in der Statistikausbildung dar. Ob diese erreicht wurden, mit Hilfe welcher Komponenten und mit welchen Formen des Einsatzes, wurde im Rahmen der 2. Durchführung der Evaluationsstudie empirisch überprüft. Die erwünschten Fähigkeiten stellen somit einen erwünschten Sollzustand der Studierenden als Output der Statistikausbildung dar.

Im Rahmen dieser Arbeit wurden acht zentrale Fähigkeiten herausgearbeitet und nach dem folgenden Schema diskutiert:

1. Es erfolgt eine exakte Beschreibung des gezielten Einsatzes der Komponenten.
2. Zudem wird begründet, warum gerade die jeweils beschriebenen Einsatzvorschläge zu den erwünschten Eigenschaften führen sollten.
3. Außerdem wird diskutiert, mit Hilfe welcher Methoden an welchen Variablen überprüft werden soll, ob die Komponenten tatsächlich zur Zielerreichung führen (=Überprüfung der internen Validität).

#### Die acht zentralen Fähigkeiten:

- Verständnis für statistische Methoden - Methodenverständnis
- Verständnis für gezielten und richtigen Einsatz statistischer Methoden zur Problemlösung - Anwendungsverständnis
- Verständnis für Zusammenhänge und Abhängigkeiten zwischen den Methoden
- Verständnis für mathematische Konzepte und schrittweise Beweisführung/Herleitung
- Fähigkeit, statistische Reports zu verfassen zur Dokumentation des statistischen Arbeitens
- Souveräner Umgang mit neuen Medien und Verständnis für die notwendige Zusammengehörigkeit von "Computational" und "Statistics"
- Fähigkeit, organisiert und selektiv zu lernen, Prioritäten beim Lernen zu setzen
- Erster Überblick über Anwendungsfelder, Relevanz und Rolle der Statistik in der Realität

Die vollständige Diskussion dieser acht Fähigkeiten befindet sich im Anhang J. Der Leser, der die Lernsoftware und das Einsatzkonzept nutzen möchte, kann sich dort informieren. Betrachten wir die Diskussion der ersten Fähigkeit an dieser Stelle exemplarisch:

#### Verständnis für statistische Methoden - Methodenverständnis

⇒ Theoriekomponenten, Skript/Vorlesungsfolien, Glossar, Statistiklabor, Minute Papers

1. Die Theoriekomponenten werden in der Vorlesung zur Vermittlung statistischer Methoden eingesetzt. Sie werden ergänzt durch Vorlesungsfolien/Skript, die die Formalia der statistischen Methoden zum Teil ausführlicher darstellen. Zum einen ist es notwendig, die Theoriekomponenten zu ergänzen, da diese oft sehr knapp gefasst sind. Zum anderen können die Ergänzungen als Individualisierung betrachtet werden. Dozenten ergänzen Inhalte, die ihnen persönlich wichtig sind, um bestimmte statistische Inhalte zu vermitteln. Des Weiteren dienen die Theoriekomponenten der Nachbereitung der Vorlesungen durch

die Studierenden. Auch Inhalte, wie zum Beispiel die Erhebungsverfahren, die in der Vorlesung nur kurz angesprochen werden, können sich die Studierenden eigenständig durch die Theoriekomponenten aneignen. Dies gilt bei Interesse auch für weiterführende statistische Methoden.

Das Glossar wird sowohl in der Vorlesung als auch in der Nachbereitung eingesetzt. Statistische Methoden oder Teile werden knapp, vergleichbar mit Lexikoneinträgen, beschrieben. Vermittelt die Theoriekomponente die Klassierung einer Häufigkeitstabelle, kann im Glossar *Häufigkeitstabelle klassiert* nachgeschlagen werden. Die im Glossar kurz beschriebenen Begriffe sind untereinander verlinkt. Es entsteht ein Überblick darüber, welche Begriffe bei bestimmten statistischen Methoden eine Rolle spielen. Die praktische Umsetzung und Anwendung der statistischen Methoden erfolgt im Statistiklabor. So kann z.B. die klassierte Häufigkeitstabelle grafisch durch das Histogramm abgebildet werden. Man kann unterschiedliche Klassenbreiten wählen und deren Auswirkung auf die grafische Darstellung der Dichte der Daten illustrieren. Dazu kann der Grafik-Wizard eingesetzt werden, wenn der Studierende sich allein auf die Ideen konzentrieren soll. Es besteht außerdem die Möglichkeit, den R-Kalkulator einzuführen und durch die Anwendung von Bibliotheken erste Kontakte mit Programmiersprachen herzustellen.

Anhand von Minute Papers wird überprüft, ob die vermittelte statistische Methode verstanden wurde.

2. Statistisches Verständnis soll besonders durch die Kombination der beschriebenen Komponenten erreicht werden. Jede Komponente erfüllt eine andere Aufgabe, zusammen sollen sie statistisches Methodenverständnis erzeugen. Das theoretische Verständnis für eine Methode wird durch die Theoriekomponente, das Glossar und Skript geschaffen. Die konkrete Anwendung wird im Labor gezeigt. In der Vorlesung ist der Studierende größtenteils passive Lernender, vertieft das dort erlernte Wissen in der Übung durch die direkte Anwendung der statistischen Methoden. Die wiederholte Anwendung der statistischen Methoden in der Übung vertieft zum einen die Idee der Methode und zum anderen deren Anwendung. Durch die Besprechung der konkreten Umsetzung in der Übung werden falsche Anwendungen und Interpretationsschwierigkeiten aufgedeckt und ihnen wird entgegengewirkt. Außerdem werden Besonderheiten der Methodenanwendung besprochen (wie z.B. Kausalzusammenhang vs. statistischer Zusammenhang).
3. In Minute Papers werden die Studierenden am Ende einer entsprechenden Vorlesung aufgefordert, Fragen zu den Ideen und Methoden zu beantworten. Minute Papers geben dem Dozenten ein Feedback. Statistisches Verständnis soll folglich als Kombination theoretischer und praktischer Aspekte vermittelt werden.  
Übungsaufgaben, deren Lösung im Labor erfolgt, überprüfen das Verständnis für die Anwendung und Interpretation der statistischen Methoden. Die Übung überprüft und vertieft dies zusätzlich und evaluiert durch direkte (teilnehmende) Beobachtung der Studierenden. Durch den Fragebogen erfolgt eine Selbsteinschätzung der Studierenden bezüglich ihres statistischen Methodenverständnisses.

#### 13.1.4 Evaluationsinstrument als Standard

**Die Idee:** Beide Durchführungen der Evaluationsstudie wurden mit je max. 40 Studierenden durchgeführt. Diese kleinen Stichprobenumfänge müssen bei der Verallgemeinerung der Evaluationsergebnisse berücksichtigt werden. Diese können zum einen aus der Qualität und dem Einsatz der Lernsoftware resultieren, zum anderen können sie allerdings eine Folge der kleinen Gruppe sein.

„Die Validität bezeichnet den Grad der Genauigkeit, mit dem die Ausprägung eines Merkmals gemessen wird. Eine Messung ist valide, wenn mit dem Messinstrument genau das gemessen wird, was der Forscher zu messen beabsichtigt und somit das verwendete Merkmal ein tatsächlicher Bestandteil des Untersuchungsgegenstandes ist.“<sup>310</sup>

<sup>310</sup>[DeckerWagner2002].



Aufgrund des geringen Stichprobenumfangs ist die externe Validität der Evaluationsergebnisse als niedrig einzustufen. Die erzielten Evaluationsergebnisse sind deshalb nicht ohne Einschränkung auf die Allgemeinheit übertragbar. Um dieser Schwäche entgegenzuwirken, wurde der Versuch unternommen, ein Evaluationsinstrument zu entwickeln. Wird das Evaluationsinstrument flächendeckend an vielen Hochschulen parallel, d.h. auf Basis hoher Stichprobenumfänge eingesetzt, steigt die Repräsentativität der Ergebnisse und somit die externe Validität. Der Aufwand der Evaluatoren sinkt durch die Benutzung des entwickelten Evaluationsinstrumentes. Das Evaluationsinstrument stellt einen ersten Schritt dar, flächendeckend auf Basis großer Stichprobenumfänge evaluieren zu können.

**Vorgehensweise und Umsetzung:** Ausgangspunkt war ein nicht zufriedenstellender Istzustand, der durch den gezielten Einsatz multimedialer sowie traditioneller Veranstaltungskomponenten in einen erwünschten Sollzustand umgewandelt werden sollte. Die dargestellten Defizite in der Statistikausbildung führten dazu, dass den Studierenden nach der Statistikausbildung wichtige statistische Kompetenzen fehlten (Istzustand). Der Sollzustand wurde beschrieben durch die Zielkompetenzen der Studierenden oder statistische Kompetenzen, die sie nach dem gezielten Einsatz bestimmter Veranstaltungskomponenten erlernt haben sollten.

Die im Folgenden dargestellte Concept Map beschreibt die Entwicklung und den Einsatz des Evaluationsinstrumentes grafisch:

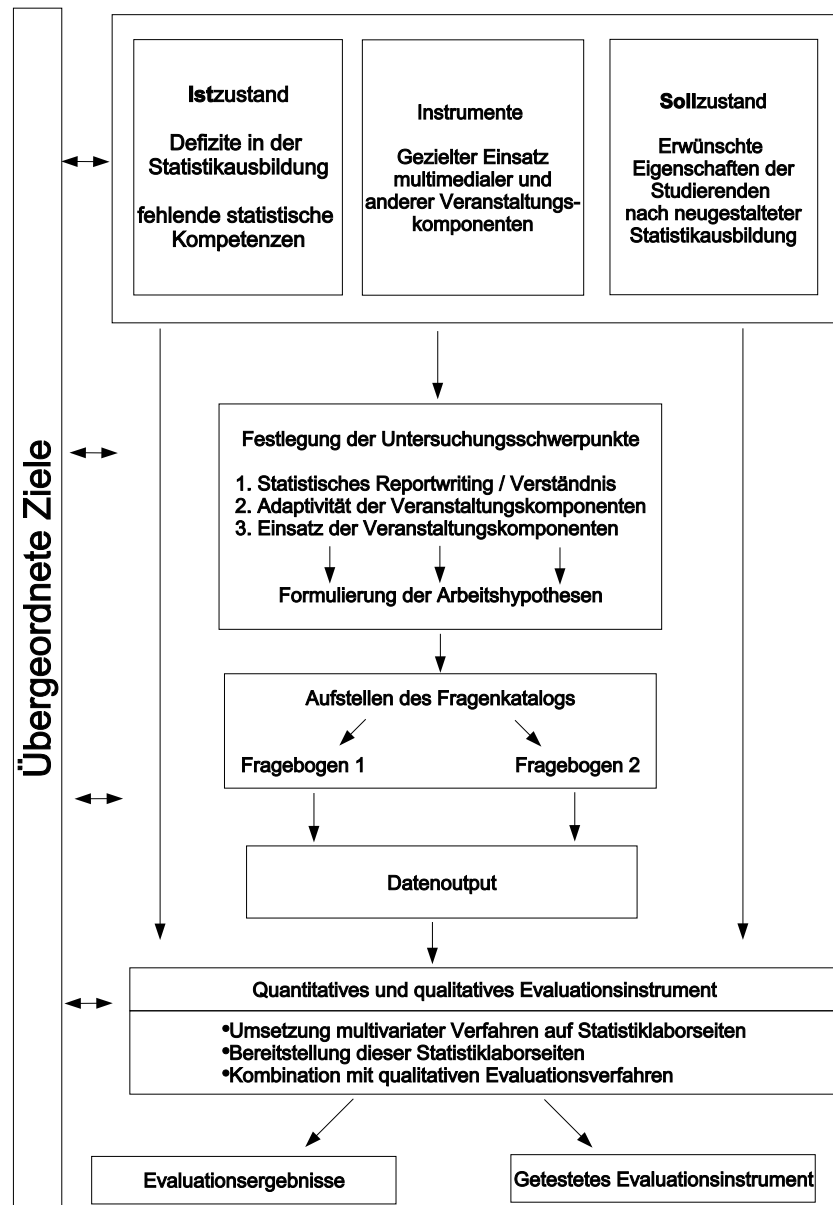


Abbildung 67: Die Entwicklung des Evaluationsinstrumentes

Auf der Diskussion der Soll/Ist-Überlegungen aufbauend wurden drei Untersuchungsschwerpunkte für die quantitative Evaluationsstudie festgelegt. Für jeden Untersuchungsschwerpunkt wurden anschließend verschiedene Arbeitshypothesen aufgestellt. In einem nächsten Schritt wurde auf Basis der Arbeitshypothesen ein Fragenkatalog entworfen, mit dessen Hilfe die interessierenden Daten bei den Studierenden erhoben wurden. Aus dem Fragenkatalog wurden zwei Fragebögen entwickelt, die den Kern der Evaluation darstellten. Durch die Art der Fragestellungen und Antwortmöglichkeiten wurde festgelegt, welche statistischen Verfahren im Rahmen der quantitativen Evaluation angewendet werden konnten.

Die mittels der Fragebögen erhobenen Daten gingen als Input in die speziell erstellten Statistiklaborseiten ein, die die Umsetzung ausgewählter statistischer Analyseverfahren ermöglichen. Diese

sind der erste Bestandteil des Evaluationsinstrumentes. Ein weiterer Bestandteil sind die der Evaluation zugrunde liegenden Evaluationsbögen sowie Arbeitshypothesen auf deren Basis die Evaluation durchgeführt wurde.<sup>311</sup>

Die Erfahrungen der 1. Durchführung zeigten, dass durch die qualitative Evaluation wichtige Evaluationsergebnisse erzielt werden können. Aus diesem Grunde wurde das Evaluationsinstrument durch qualitative Evaluationselemente ergänzt. Der zweite Evaluationsbogen enthält offene Fragen.

Als Output sind ein durch den Einsatz getestetes Evaluationsinstrument sowie Evaluationsergebnisse zum Einsatz neuer Medien in der Statistikausbildung entstanden.

## 13.2 Entwicklung und Einsatz des Evaluationsinstrumentes

### 13.2.1 Zielsetzungen und strategischer Einsatz

Während der 2. Durchführung wurden im Kern die folgenden vier Ziele verfolgt, von denen drei das Evaluationsinstrument direkt betreffen:

- Entwicklung eines Evaluationsinstrumentes, mit dessen Hilfe multimediale Lernsoftware sowie ihr Einsatz evaluiert wurden. Das Evaluationsinstrument bot den Einsatz statistischer Analyseverfahren an, konzentrierte sich somit auf die quantitative Evaluation. Die Umsetzung der statistischen Analyseverfahren fand auf speziell entwickelten Statistikkaborseiten statt.
- Test des entwickelten Evaluationsinstrumentes durch den Einsatz und die Evaluation multimedialer Lernsoftware auf Basis des Evaluationsinstrumentes. Der Test wurde am Beispiel der statistischen Lernsoftware *Statistik interaktiv komplett* durchgeführt.
- Evaluation der statistischen Lernsoftware *Statistik interaktiv komplett* und des für diese entwickelten Einsatzkonzeptes.
- Integration qualitativer Elemente<sup>312</sup> in das Evaluationsinstrument, welche die quantitativen Evaluationsmethoden ergänzten. Vorschlag einer möglichen Kombination qualitativer und quantitativer Evaluationsmethoden in Bezug auf multimediale Lernsoftware.

Während der 2. Durchführung der Evaluationsstudie wurde zu mehreren Zeitpunkten überprüft, ob die vier Ziele berücksichtigt wurden.

Betrachten wir die vierstufige Vorgehensweise zur Zielerreichung zunächst grafisch:

---

<sup>311</sup>Die Materialien sind in den Anhängen D, E und K dieser Arbeit bereitgestellt.

<sup>312</sup>Vgl. Anhang G.

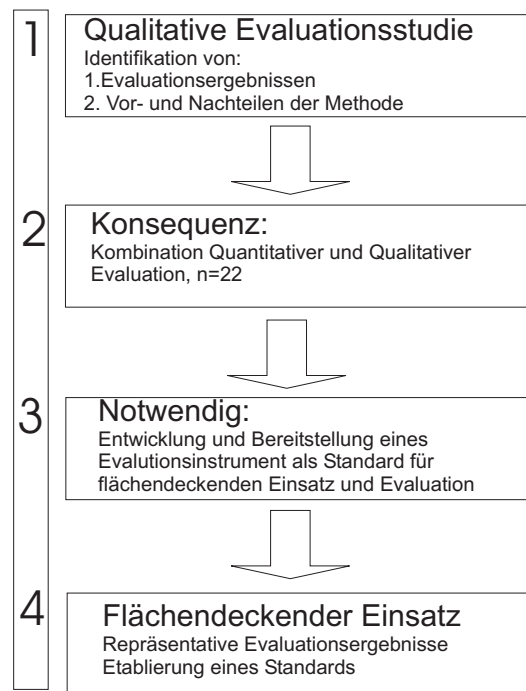


Abbildung 68: Vierstufige Vorgehensweise

1. Im Rahmen der Dissertation hat im WS 2001/2002 die 1. Durchführung der Evaluationsstudie stattgefunden. Die multimediale Lernsoftware *Statistik interaktiv komplett* (bzw. ein Prototyp) wurde zusammen mit einem Einsatzkonzept (bzw. einem Prototypen) rein qualitativ evaluiert. Die Evaluationsergebnisse wurden auf Basis eines Stichprobenumfangs von  $n=40$  erhoben. Trotz des geringen Stichprobenumfangs gestaltete sich die qualitative Evaluation als sehr aufwendig. Die 1. Durchführung diente dazu, sowohl die Stärken und Schwächen der Komponenten der multimedialen Lernsoftware als auch der des Einsatzkonzeptes zu erkennen. Die Untersuchung eines größeren Stichprobenumfangs wäre deshalb in dieser Phase nicht sinnvoll gewesen.
2. Als Konsequenz der Einschätzung und Bewertung der qualitativen Evaluationsstudie wurde nach weiteren Möglichkeiten der Evaluation gesucht und für die zweite Evaluationsstudie die quantitative Evaluation gewählt. Diese ergänzte die qualitative Evaluation und setzte besonders dort an, wo die qualitative Evaluation Mängel aufwies. Durch den geringen Stichprobenumfang wurde die Repräsentativität der Evaluationsergebnisse nicht gewährleistet. Das heißt, dass weder allgemeingültige Aussagen über die Qualität der multimedialen Lernsoftware noch die des Einsatzkonzeptes getroffen werden konnten. Aus folgenden Gründen konnte kein größerer Stichprobenumfang untersucht werden, der zu einer verbesserten Aussagefähigkeit geführt hätte und gleichzeitig mit einer angemessenen Betreuung der Teilnehmer verbunden gewesen wäre:
  - unzureichende technische Ausstattung der Universität
  - fehlendes standardisiertes Evaluationsinstrument
  - Probleme mit dem Statistikkolabor
  - hoher Aufwand bei der Betreuung der Teilnehmer
3. Die direkte Konsequenz aus diesen Überlegungen bestand in der Entwicklung und Bereitstellung eines Evaluationsinstrumentes, das einen flächendeckenden (z.B. bundesweiten) standardisierten Einsatz und Evaluation der multimedialen Lernsoftware ermöglicht.

Setzen viele Universitäten die Software auf Basis des entwickelten Einsatzkonzeptes gleichzeitig ein und evaluieren mit Hilfe des Evaluationsinstrumentes, können die erhobenen Evaluationsdaten zusammengefasst und evaluiert werden. Die Bereitstellung eines Evaluationsinstrumentes standardisiert somit den flächendeckenden Einsatz/Evaluation. Außerdem erleichtert es die Arbeit und somit die Bereitschaft der Dozenten, teilzunehmen. Das Evaluationsinstrument wurde flexibel gestaltet. Die Dozenten sollten nicht zu starken Restriktionen unterworfen werden. Die Schwierigkeit bestand darin, dass Flexibilität in einem gewissen Widerspruch zur Schaffung eines Standards steht.

### 13.2.2 Untersuchungsschwerpunkte, Arbeitshypothesen, Evaluationsbögen

Die Defizite und daraus abgeleiteten Zielkompetenzen des Studierenden führen zu den folgenden Untersuchungsschwerpunkten:

1. statistisches Reportverfassen, statistisches Verständnis, Methodenverständnis
2. Adaptivität multimedialer Komponenten
3. Gestaltung des Einsatzkonzeptes

Die Formulierung der Arbeitshypothesen erfolgte auf Basis der formulierten Zielkompetenzen der Studierenden sowie der drei gebildeten Untersuchungsschwerpunkte. Sie bildeten die Grundlage für die im Rahmen der Evaluation zu erhebenden Daten.

Jede Arbeitshypothese wurde entweder qualitativ oder quantitativ untersucht. Die Arbeit enthält die Arbeitshypothesen sowie die Argumentation der Auswahl statistischer Verfahren. Dabei wird in zwei Schritten vorgegangen. Zuerst wird beschrieben, wie evaluiert wurde: unter Anwendung welcher Evaluationsverfahren. Im zweiten Schritt wird dies begründet, der Frage nach dem Warum nachgegangen. Die vollständige Diskussion steht dem Leser im Anhang K<sup>313</sup> bereit. Auf der folgenden Seite erfolgt die Darstellung aller Arbeitshypothesen sowie die exemplarische Diskussion einer Hypothese in den beschriebenen zwei Schritten. Die Arbeitshypothesen sind in vier Bereiche gegliedert:

1. statistisches Reportverfassen, statistisches Verständnis, Methodenverständnis
2. Adaptivität (Nutzungsgrad-Nutzungsdauer, Akzeptanz, Interaktivität) der multimedialen Komponenten
3. Einsatzkonzept für multimediale Veranstaltung
4. Lernmanagementsystem Blackboard (BB)

- **Statistisches Reportverfassen, Statistisches Verständnis, Methodenverständnis**

$H_1$  : Die Vermittlung der Ideen der mathematischen Seite der statistischen Methoden ist durch die neuen Medien besser möglich.

$H_2$  : Die Idee der Erstellung statistischer Reports kann durch die neuen Medien besser vermittelt werden.

$H_3$  : Die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik wird durch die Multimediasoftware *Statistik interaktiv komplett* während des Semesters signifikant verändert.

$H_4$  : Den Studierenden wird vermittelt, dass es sich bei der Statistik um eine fächerübergreifende Disziplin handelt, die in vielen Fächern eine wichtige Rolle spielt.

$H_5$  : Statistisches Reportverfassen kann im Team besser erlernt werden als in Einzelarbeit.

$H_6$  : Die Fähigkeit statistische Reports zu verfassen, wird durch den kontinuierlichen Einsatz des Statistiklabors in der Vorlesung und Übung verbessert.

---

<sup>313</sup>Dort ist gekennzeichnet, welche Fragen der Evaluationsbögen den jeweiligen Arbeitshypothesen zugeordnet sind.

$H_7$  : Die Fähigkeit, statistische Reports zu verfassen, wird durch die Lösung von Aufgaben im Statistiklabor verbessert.

$H_8$  : Die Fähigkeit, statistische Reports zu verfassen, wird durch die Erstellung von Minireports innerhalb der ersten Wochen des Semesters verbessert.

$H_9$  : Die Fähigkeit, statistische Reports zu verfassen, wird durch die dreimal im Semester stattfindende bewertete Reporterstellung verbessert.

- **Adaptivität (Nutzungsgrad-Nutzungsdauer, Akzeptanz, Interaktivität) der multimedialen Komponenten**

$H_{10}$  : Es besteht ein Zusammenhang zwischen verschiedenen Komponenten des Einsatzkonzeptes der multimedialen Statistikveranstaltung bezüglich der Eigenaktivität der Studierenden.

$H_{11}$  : Der gezielte Einsatz des Laborobjektes *Kalkulator* ermöglicht eine schrittweise Heranführung an Programmiersprachen zur Verringerung von Barrieren.

- **Einsatzkonzept für multimediale Veranstaltung**

$H_{12}$  : Die Studierenden erwarten von einer multimedialen Statistikveranstaltung, dass die eingesetzten Komponenten (Lernmanagementsystem, Lernsoftware, Statistiklabor) für sie in naher Zukunft (Studium und Beruf) Relevanz haben.

$H_{13}$  : Die Studierenden erwarten, dass ihnen das Erlernen statistischer Inhalte durch die multimediale Lernsoftware erleichtert wird.

$H_{14}$  : Das gemeinsame Erstellen von Statistiklaborseiten in der Vorlesung und Übung zeigt die Möglichkeiten und Arbeitserleichterungen auf und animiert zur Nutzung.

- **Lernmanagementsystem Blackboard (BB)**

$H_{15}$  : Die Organisation des Lernens und der Umgang mit neuen Medien wird durch deren Einsatz in einem Lernmanagementsystem verbessert.

$H_{16}$  : Traditionelle und multimediale Komponenten lassen sich in einem Lernmanagementsystem sinnvoll miteinander kombinieren.

$H_{17}$  : Die Teamarbeit der Studierenden wird durch das Lernmanagementsystem positiv unterstützt.

Wir diskutieren exemplarisch die Hypothese 3:

- Es erfolgt eine Einstellungsmessung über das Rosenberg-Modell, d.h. es wird gemessen, ob und wie sich die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik im Laufe des Semesters verändert hat. Über einen zweiseitigen T-Test wird überprüft, ob eine signifikante Veränderung stattgefunden hat.
- Das Ziel besteht darin, die Einstellungsveränderung der Studierenden zu überprüfen und festzustellen, durch welche getroffenen Maßnahmen diese bewirkt wurden. Hat z.B. das Reportverfassen zu einer positiven Einstellungsveränderung bei den Studierenden geführt? Wird im Mittel durch den T-Test eine signifikante Veränderung bestätigt, werden in nächsten Schritten mögliche Korrelationen mit anderen Fragen der Fragebögen untersucht. Auf diese Weise kann bestimmt werden, durch welche Maßnahmen die Einstellung wie verändert wurde, d.h. es findet eine Überprüfung der Wirksamkeit der Maßnahmen statt - auf Basis der Einstellung der Studierenden.

Das Ziel bestand darin, auf Basis der Arbeitshypothesen Fragen und Antwortkategorien zu entwickeln, die die Anwendung statistischer Analyseverfahren ermöglichten. Ergaben sich an einigen Stellen durch die Anwendung anderweitiger Methoden (z.B. aus der Marktforschung) Vorteile, so wurden diese angewendet.

Es handelte sich um schriftliche Befragungen, die durch Rückfragen begleitet wurden, d.h. die Studierenden konnten bei Verständnisschwierigkeiten Rücksprache mit den Dozenten halten. Da die

2. Durchführung der Evaluationsstudie als Pretest betrachtet wurde, wurden die Fragebögen hinsichtlich ihrer Gestaltung vor ihrem Einsatz nicht experimentell überprüft. Die Fragebögen wurden in der 2. Durchführung als 1. und 2. Evaluationsbogen<sup>314</sup> bezeichnet.

### 13.2.3 Der Evaluationsreport

Der Entwicklungsprozess des Evaluationsinstrumentes bestand aus den vier aufeinander folgenden Schritten:

1. Konzeption der Arbeitshypothesen und Aufstellung entsprechender Fragebögen
2. Datenerhebung während und nach der Durchführung der Evaluationsstudie (Pretest), kritische Betrachtung der Arbeitshypothesen und ihre daraus resultierende Eliminierung und Neugenerierung
3. Konzeption und Entwicklung entsprechender Statistiklaborseiten zur Auswertung des erhobenen Datenmaterials
4. Entwicklung des Evaluationsinstrumentes durch Kombination der Statistiklaborseiten mit qualitativen Evaluationsverfahren. Dokumentation im Evaluationsreport

Es erfolgt die Bereitstellung des Evaluationsreports. Dieser besteht aus den Laborgrafiken (Output der entwickelten Statistiklaborseiten<sup>315</sup>), sowie Interpretations- und Auswertungsmodulen inferenzstatistischer Methoden. Um das Verfassen eines Evaluationsreports zu demonstrieren, werden exemplarisch konkrete Interpretationen vorgenommen. Dies erfolgt auf Basis des im Rahmen der Evaluationsstudie erhobenen Datenmaterials. Das Ziel der Bereitstellung des Evaluationsreports besteht darin, den Umgang mit dem entwickelten Evaluationsinstrument zu demonstrieren.

Der entstandene Evaluationsreport besteht aus zwei Teilen:

- Dokumentation der quantitativen Evaluation
- Dokumentation der qualitativen Evaluation

Zur Erstellung des Reports wurde die Berichterstellungsfunktion eingesetzt. Der entstehende statistische Report hat mehrere Funktionen:

- Er stellt dem Leser die Idee des statistischen Reports vor,
- zeigt ihm, welche Umsetzungsmöglichkeiten in der entwickelten Lernsoftware *Statistik interaktiv komplett* geschaffen wurden,
- wertet die Evaluationsstudie exemplarisch aus und
- zeigt, wie quantitative und qualitative Evaluationsmethoden kombiniert werden können.

Der Report beschreibt einerseits, wie ein Evaluationsreport gestaltet sein könnte, d.h. er hat den Charakter eines Metareports, andererseits enthält er konkrete Ausführungen, die einzelne Laborseiten des Evaluationsinstrumentes demonstrieren.<sup>316</sup>

---

<sup>314</sup>Vgl. Anhang D und E.

<sup>315</sup>Alle Statistiklaborseiten des Evaluationsinstrumentes stehen auf der CD-ROM im Anhang A dieser Arbeit bereit.

<sup>316</sup>Im anschließend dargestellten Report wird auf Laborseiten verwiesen. Diese enthalten alle die Endung spf. Diese wird nicht einzeln ausgewiesen, um den Lesefluss nicht zu beeinträchtigen.

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

## **Evaluationsbericht in Form eines statistischen Reports**

Der folgende statistische Report setzt sich zusammen aus:

1. Problembeschreibung
2. Auswahl statistischer Methoden
3. Anwendung der gewählten Methoden auf das Datenmaterial und Interpretation der Datenanalyseergebnisse
4. Qualitative Befragung der Studierenden in schriftlicher Form. Zusammenfassung der Antworten.
5. Rückbezug auf die Realität

### **1. Problembeschreibung**

Im Rahmen des Forschungsprojektes NEUE STATISTIK wurde an der Universität Bielefeld eine Evaluationsstudie durchgeführt. Diese hatte das Ziel, die entwickelte multimediale Lernsoftware *Statistik interaktiv komplett* zu evaluieren. Die Komponenten dieser wurden im Rahmen eines speziell entwickelten Einsatzkonzeptes überprüft. Die Qualität dieses Einsatzkonzeptes wurde ebenfalls evaluiert. Die Evaluationsstudie kann als Pretest des entwickelten Evaluationsinstrumentes angesehen werden. Die erzielten Evaluationsergebnisse basieren auf einem Stichprobenumfang von  $n=20$ , sind daher nicht repräsentativ. Das Evaluationsinstrument stellt eine Möglichkeit dar, eine Basis für eine flächendeckende, bundesweite Evaluation zu schaffen. Das übergeordnete Ziel der Evaluation bestand darin, Schwächen und Stärken der Lernsoftware und des Einsatzkonzeptes zu identifizieren und eine schrittweise Optimierung beider anzustreben. Der vorliegende Evaluationsbericht hat zwei Aufgaben: die Demonstration und die Anwendung des entwickelten Evaluationsinstrumentes. Die Demonstration erfolgt durch die konkrete Anwendung des Instrumentes. Dabei wird exemplarisch vorgegangen.

### **2. Auswahl statistischer Methoden**

Die Datenerhebung erfolgte auf Basis zweier Evaluationsbögen, die zu Semesterbeginn und –ende durch die Teilnehmer ausgefüllt wurden. Die Ziele der Evaluationsstudie bestimmten die gestellten Fragen und vorgegebenen Antwortmöglichkeiten. Die angewendeten statistischen Methoden determinierten die Gestaltung der Fragen und Antwortmöglichkeiten. Ein Beispiel zur Verdeutlichung dieser Beziehungen: besteht das Ziel darin, deskriptiv Informationen über die technische Ausstattung der Teilnehmer zu gewinnen,



Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

können z.B. die relativen Häufigkeiten dargestellt werden (statistische Methode). Um relative Häufigkeiten bestimmen zu können, werden binäre Antwortmöglichkeiten benötigt. Es wird deutlich, dass das Ziel den Rahmen für die Methode und die vorzuziehenden Antwortmöglichkeiten festlegt.

Größtenteils werden deskriptive Analyseverfahren angewendet, um sich einen Überblick über das Datenmaterial zu verschaffen. Um auch Zusammenhänge einzelner Variablen/Merkmale zu untersuchen, werden Kontingenztabellen aufgestellt und Regressionsanalysen durchgeführt. Einige auf Basis grafischer Darstellungen gewonnenen Evaluationsergebnisse werden mit Hilfe des T-Tests überprüft.

### 3. Anwendung der gewählten Methoden auf das Datenmaterial Interpretation der Datenanalyseergebnisse

Hier wird unterteilt nach den folgenden Gebieten:

- a) Auswertung deskriptiver Fragestellungen - Typisierung
- b) Evaluation des statistischen Reportwritings, des statistischen Verständnisses, des Methodenverständnisses
- c) Evaluation der Adaptivität der multimedialen Komponenten
- d) Evaluation des entwickelten Einsatzkonzeptes
- e) Evaluation des eingesetzten Lernmanagementsystems Blackboard

Der Report enthält zu jedem Gebiet allgemeine Informationen über die einzelnen Laborseiten. Exemplarisch werden einzelne Laborseiten konkret angewendet, um die Anwendung und Möglichkeiten des Evaluationsinstrumentes zu demonstrieren. Alle Laborseiten haben die Endung .spf.

#### 3.a) Auswertung deskriptiver Fragestellungen – Typisierung

Über die Auswertung deskriptiver Fragestellungen wird ein erster Eindruck von den Teilnehmern einer multimedialen Statistikveranstaltung gewonnen. Dies geschieht auf den Laborseiten **Deskriptiv1-8** des Evaluationsinstrumentes.

Durch verschiedene Stabdiagramme wird ein Überblick über die folgenden Aspekte gewonnen:

- die technische Ausstattung der Teilnehmer (Laborseite: **Deskriptiv1**)
- die Internetnutzung der Teilnehmer (Laborseite: **Deskriptiv2**)
- die Charakterisierung des Fachs Statistik durch die Teilnehmer (Laborseiten: **Deskriptiv3+4**)
- die Charakterisierung und Bewertung von *Statistik interaktiv komplett* (Laborseiten: **Deskriptiv5-8**)

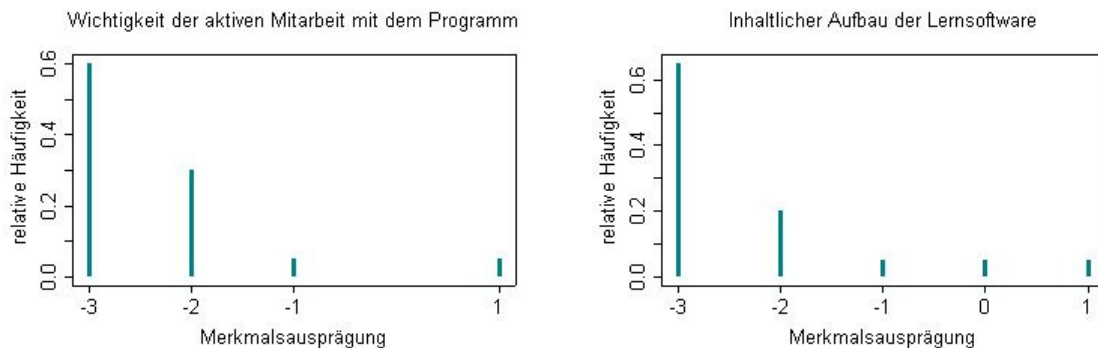
Mit Hilfe den folgenden Grafiken wird untersucht, wie die Studierenden die eingesetzte Lernsoftware *Statistik interaktiv komplett* bewerten und charakterisieren. Dabei werden verschiedene Aspekte untersucht, die den Überschriften der Grafiken zu entnehmen sind. Auf den folgenden Grafiken wird untersucht, wie

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität Bielefeld

wichtig den Studierenden die verschiedene Aspekte des Lernprogramms sind. Zwei werden exemplarisch betrachtet:

(Laborseite: **Deskriptiv5**):

(Legende:  
-3: sehr wichtig  
bis  
3: sehr unwichtig)



Die konkrete Anwendung des Evaluationsinstrumentes hat zu dem Ergebnis geführt, dass die Studierenden sowohl den inhaltlichen Aufbau als auch die aktive Arbeit mit *Statistik interaktiv komplett* als sehr wichtig bewerten.

### 3.b) Evaluation des statistischen Reportwritings, des statistischen Verständnisses, des Methodenverständnisses

Nach den ersten deskriptiven Analysen werden die aufgestellten Arbeitshypothesen überprüft. Dabei werden zunächst die untersucht, die quantitativ im Statistiklabor ausgewertet werden können. An dieser Stelle wird exemplarisch die Hypothese 3: *Die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik wird durch die Multimediasoftware Statistik interaktiv komplett während des Semesters signifikant verändert.* betrachtet. Mit Hilfe des Rosenberg-Modells (vgl. S. 200ff. in Kroeber-Riel, Weinberg Konsumentenverhalten, 8. Auflage, Verlag Franz Vahlen München 2003) wird untersucht, ob und wie sich die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik im Laufe des Semesters verändert hat. Die folgenden Berechnungen werden auf der Laborseite **Hypothese3** durchgeführt:

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

Welch Two Sample t-test

```
data: summe1 and summe2
t = 0.7026, df = 36.752, p-value = 0.4867
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.507699 3.107699
sample estimates:
mean of x mean of y
4.55 3.75
```

Die dem T-Test zugrundeliegende Hypothese besagt, dass die Mittelwerte der gemessenen Einstellungen der Studierenden zu Semesterbeginn und Semesterende gleich sind. Diese kann nicht abgelehnt werden, da die Überschreitungswahrscheinlichkeit  $p=0.4867$  beträgt. D.h. es findet keine signifikante Veränderung der Einstellungen der Studierenden in Bezug auf die Realitätsnähe der Statistik zu einem Signifikanzniveau von 5% oder 10% statt.

Anschließend erfolgt eine deskriptive Untersuchung dieser Hypothese auf der Laborseite **Hypothese4**. Mit Hilfe eines Chi-Quadrat-Tests wird auf dieser außerdem untersucht, ob diese Veränderung signifikant ist. Die Einschätzung der Studierenden bezüglich Relevanz statistischer Methoden in verschiedenen wirtschaftlichen Bereichen wird auf der Laborseite **Hypothese4b** grafisch dargestellt.

Auf den Laborseiten **Hypothese5+5b** wird überprüft, ob die Studierenden die Meinung vertreten, dass statistisches Reportwriting im Team besser erlernt werden kann, als in Einzelarbeit. Außerdem wird die im Einzelkonzept implementierte schrittweise Vermittlung des statistischen Reportwritings über:

1. Aufgabenlösungen im Statistiklabor,
2. Verfassen von Minireports,
3. Erstellen von drei Reports im Verlauf des Semesters

auf seine Wirksamkeit hin überprüft.

Im nächsten Schritt wird der Zusammenhang zwischen der Vorbereitung:

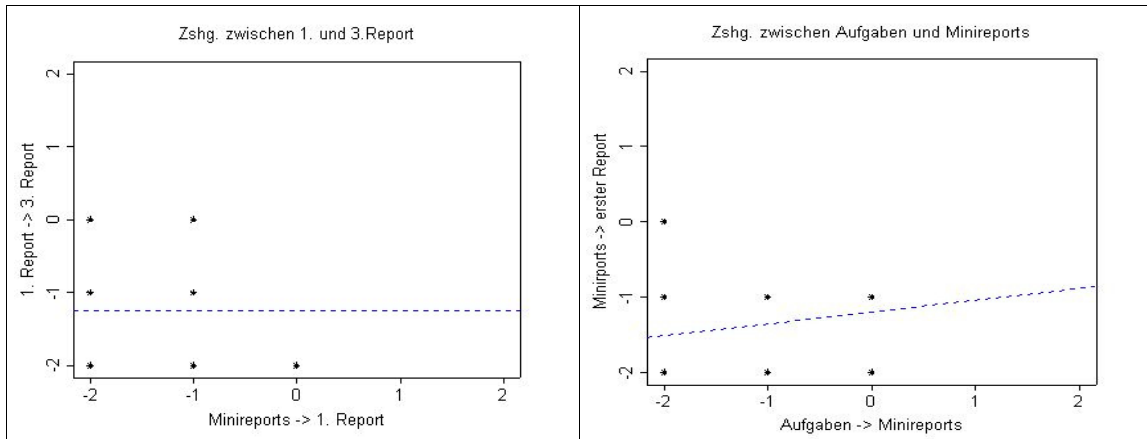
- der Aufgaben auf die Minireports und
- der Minireports auf den ersten Report

durch Anwendung der Regressionsanalyse untersucht. Alle Daten basieren auf den Einschätzungen der Studierenden. Außerdem wird der Zusammenhang zwischen:

- der Bearbeitung des ersten Reports und
- der Vorbereitung auf den dritten Report

durch Anwendung der Regressionsanalyse untersucht. Alle Daten basieren auf den Einschätzungen der Studierenden. Betrachten wir die konkrete Anwendung:

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld



Die konkrete Anwendung kommt zu dem folgenden Ergebnis: Es besteht kein linearer Zusammenhang zwischen der Anfertigung der Minireports auf den ersten Report, sowie kein linearer Zusammenhang zwischen der Anfertigung der Aufgaben auf die Minireports. Die berechneten Korrelationskoeffizienten von Bravais-Pearson, die beide nah bei Null liegen, unterstützen diese Ergebnisse:

Der Korrelationskoeffizient von Bravais-Pearson ist:

0,1921168 (linke Grafik)

Der Korrelationskoeffizient von Bravais-Pearson ist:

0 (rechte Grafik)

An dieser Stelle werden die Grenzen der quantitativen Evaluation deutlich. Es erscheint angebracht, die subjektiven Einschätzungen der Studierenden durch die in den Reports erzielten Ergebnisse zu überprüfen.

### 3.c) Evaluation der Adaptivität der multimedialen Komponenten

Überprüfung des Einsatzkonzeptes der multimedialen Statistikveranstaltung:

In diesem Bereich wird untersucht, ob ein Zusammenhang zwischen verschiedenen Komponenten des Einsatzkonzeptes in Bezug auf die Eigenaktivität der Studierenden besteht. Die Eigenaktivität wird deskriptiv bezüglich ausgewählter Aspekte auf der Laborseite **Hypothese10a** untersucht. Die Laborseite **Hypothese10b** stellt dar, in welchen Bereichen außerhalb der Vorlesung die Studierenden das Statistiklabor anwenden.

Im Anschluss werden mögliche Zusammenhänge zwischen diesen Merkmalen in einer mehrdimensionalen Kontingenztafel untersucht (0:nein, 1:ja in Bezug auf Eigenaktivität). In der konkreten Anwendung ist die

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

Kontingenztafel wie folgt gestaltet (Laborseite: **Hypothese10c**):

		Übungsaufgaben multimediale- Komp	0	0	1	1	
			0	1	0	1	
Beteiligung	Laborobjekte						
0	0		0	0,05	0	0	0,05
0	1		0	0	0	0,05	0,05
1	0		0,25	0,10	0,10	0,15	0,60
1	1		0,15	0,05	0,05	0,05	0,30
			0,40	0,20	0,15	0,25	1,00

Es sind die Zusammenhänge der Merkmale

- Lösung zusätzlicher Übungsaufgaben im Statistiklabor,
- Betrachtung zusätzlicher multimedialer Komponenten,
- Aktive Beteiligung in Vorlesung und/oder Übung,
- Nutzung zusätzlicher Laborobjekte

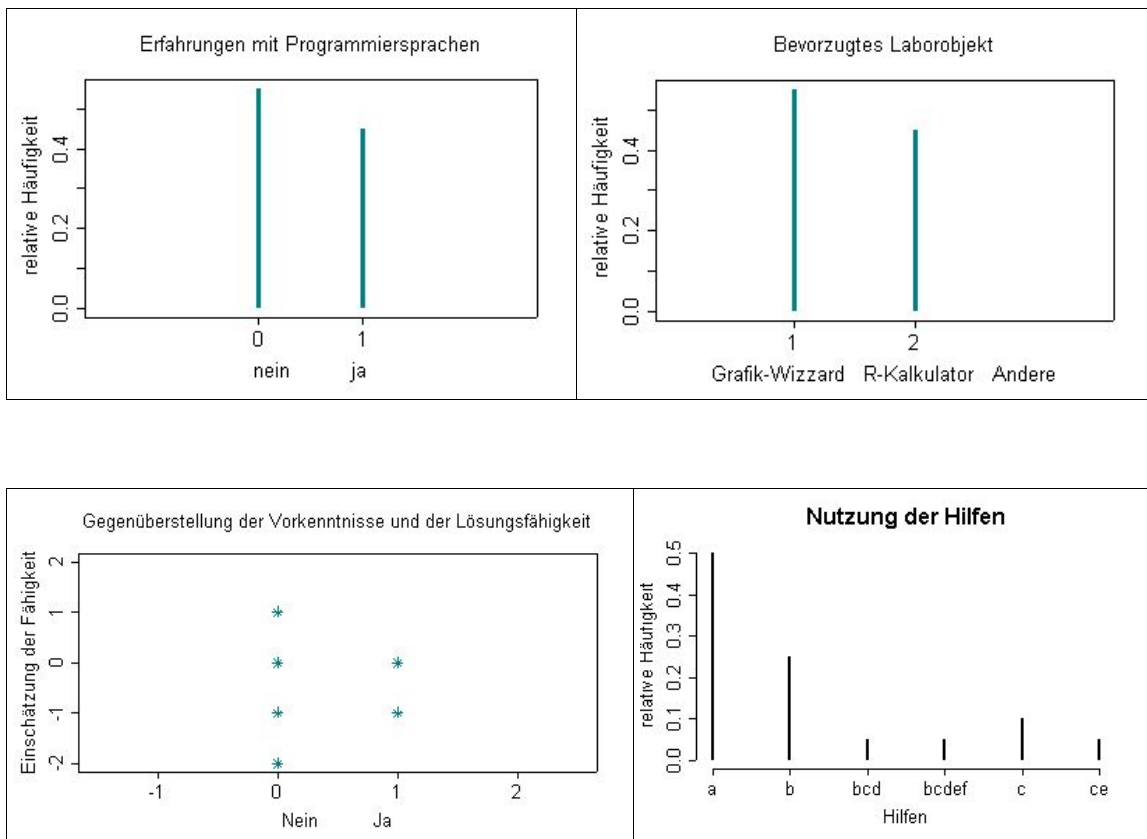
in der Kontingenztafel abgebildet. In den Zellen befinden sich die relativen Häufigkeiten der gegebenen Antworten der Studierenden. Interpretieren wir ausgewählte Zellen der Kontingenztafel exemplarisch:

25% der Teilnehmer zeigen nur Eigenaktivität bezüglich der Beteiligung an Vorlesung oder Übung und keine Beteiligung in den anderen drei Bereichen. 30 % (Randhäufigkeit) der Teilnehmer sind sowohl in der Vorlesung und Übung, als auch bezüglich der Nutzung zusätzlicher Laborobjekte aktiv. 20 % (Randhäufigkeit) der Teilnehmer zeigen keine Aktivität bezüglich der Lösung zusätzlicher Übungsaufgaben, aber Aktivität bezüglich der Betrachtung zusätzlicher multimedialer Komponenten.

In einem nächsten Schritt wird untersucht, ob der gezielte Einsatz des R-Kalkulators eine schrittweise Heranführung der Studierenden an Programmiersprachen ermöglicht und somit bestehende Barrieren verringert. Zunächst wird deskriptiv untersucht ob (Laborseite: **Hypothese11a**):

- Programmierkenntnisse bestehen,
  - ein kleines Problem mit Hilfe des Statistiklabors gelöst werden würde,
  - der R-Kalkulator dem Grafik-Wizzard vorgezogen wird, wie hoch der R-Kalkulator bezüglich statistischen Problemlösens bewertet wird,
  - wie wichtig die Erlernung seiner Anwendung eingeschätzt wird
- und
- ob und für welche Zwecke bereitgestellte Hilfen (Online-Tutorial) zum Statistiklabor genutzt wurden

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität Bielefeld



Legende für das untere Stabdiagramm:

- a: nein
- b: R-Kalkulator
- c: Grafik-Wizzard
- d: Datensatzimport
- e: R-Grafik
- f: Datensatzexport

Anhand der oberen beiden Stabdiagramme wird deutlich, dass sich die Teilnehmer in zwei ungefähr gleich große Gruppen teilen: sowohl in Bezug auf Erfahrungen mit Programmiersprachen, als auch in Bezug auf die Präferenzen bezüglich des R-Kalkulators bzw. des Grafik-Wizzards.

Bei der Gegenüberstellung der Vorkenntnisse bezüglich Programmiersprachen und der Einschätzung ein gegebenes Problem mit dem Statistiklabor zu lösen, ist zu erkennen, dass bei den Personen ohne Vorkenntnisse ein größerer Range als bei den Personen mit Vorkenntnissen vorhanden ist. Interessant hierbei ist, dass Personen ohne Vorkenntnisse die beste Einschätzung abgegeben haben. Dies veranlasst zu der Annahme,

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

dass das Statistiklabor auch für Personen ohne Programmierkenntnisse sehr gut geeignet ist.

Das letzte Stabdiagramm zeigt, dass die bereitgestellten Hilfen zum Statistiklabor fast gar nicht genutzt werden. Eine Überprüfung auf Basis eines größeren Stichprobenumfangs scheint angebracht, um den Stellenwert des aufwendig zu entwickelnden Online-Tutorials festzustellen.

Abschließend wird auf der Laborseite **Hypothese11b** mit Hilfe des zweiseitigen T-Tests überprüft, ob im Verlauf des Semesters eine signifikante Veränderung bezüglich der Anwendung des R-Kalkulators stattgefunden hat:

Vergleich: Einschätzung der Studierenden, dass der R-Kalkulator ein elementares Mittel zur statistischen Problemlösung darstellt.

Welch Two Sample t-test

```
data: problem1 and problem2
t = 0.2456, df = 33.393, p-value = 0.8075
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.3640142 0.4640142
sample estimates:
mean of x mean of y
-1.40 -1.45
```

Die dem T-Test zugrundeliegende Hypothese besagt, dass die Mittelwerte der Bewertungen der Studierenden zu Semesterbeginn und Semesterende gleich sind. Diese kann nicht abgelehnt werden, da die Überschreitungswahrscheinlichkeit  $p=0.8075$  beträgt. D.h. es findet keine signifikante Veränderung der Bewertungen der Studierenden in Bezug auf die Anwendung des R-Kalkulators zu einem Signifikanzniveau von 5% oder 10% statt.

Anschließend wird auf der Laborseite **Hypothese11b** auch mit Hilfe des zweiseitigen T-Tests überprüft, ob eine signifikante Veränderung im Verlauf des Semesters bezüglich der Antwort auf die Frage: „Wie wichtig erachten Sie die Erlernung und Anwendung des R-Kalkulators?“ stattgefunden hat.

Wichtigkeit der Erlernung des R-Kalkulators;  
Einschätzung durch die Studenten.

Welch Two Sample t-test

```
data: erlernung1 and erlernung2
t = 1.5297, df = 33.436, p-value = 0.1355
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.08234342 0.58234342
sample estimates:
mean of x mean of y
-1.55 -1.80
```

Die dem T-Test zugrundeliegende Hypothese besagt, dass die Mittelwerte der Bewertungen der Studierenden zu Semesterbeginn und Semesterende gleich sind. Diese kann nicht abgelehnt werden, da die Überschreitungswahrscheinlichkeit  $p=0.1355$  beträgt.

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

D.h. es findet keine signifikante Veränderung der Bewertungen der Studierenden in Bezug auf Wichtigkeit R-Kalkulators zu einem Signifikanzniveau von 5% oder 10% statt.

Auch bezüglich dieser beiden Testergebnisse erscheint es sinnvoll, das Ergebnis auf Basis eines großen Stichprobenumfang zu überprüfen. Das Ergebnis muss allerdings nicht durch den geringen Stichprobenumfang zustande gekommen sein. An dieser Stelle ist es empfehlenswert, die quantitative durch die qualitative Evaluation zu ergänzen.

### 3.d) Evaluation des entwickelten Einsatzkonzeptes

In diesem Bereich wird überprüft, wie die Studierenden

- den Einfluss des Statistiklabors auf ihre später notwendigen statistisch-methodischen Fähigkeiten,
- die Wichtigkeit des Einsatzes des Lernmanagementsystems Blackboard und
- die Notwendigkeit von *Statistik interaktiv komplett* zum Erlernen der statistischen Vorlesungsinhalte

einschätzen und/oder bewerten. Zunächst werden die Antworten der Studierenden bezüglich der drei genannten Aspekte auf der Laborseite **Hypothese12** deskriptiv dargestellt. Auf der Laborseite **Hypothese12a** wird die Bewertung der Studierenden bezüglich des Lernmanagementsystems und der Lernsoftware grafisch dargestellt. Mit Hilfe einer Kontingenztafel wird der Zusammenhang beider Merkmale untersucht. Anschließend wird auf der Laborseite **Hypothese12b** mit Hilfe der Regressionsanalyse der Zusammenhang der beiden Merkmale *Einfluss des Statistiklabors* und *Wichtigkeit von Blackboard* analysiert. Das Ergebnis wird durch die Berechnung des Korrelationskoeffizienten von Bravais numerisch untermauert.

Es erfolgt eine weitere Analyse im Rahmen der Evaluation der im Einsatzkonzept getroffenen Maßnahmen: die Einstellung/Erwartung der Studierenden bezüglich der Erleichterung des Erlernens statistischer Inhalte durch die multimediale Lernsoftware wird mit Hilfe des Rosenberg-Modells gemessen (vgl. Laborseite: **Hypothese13**).

Es wird zudem untersucht, ob sich diese im Verlauf des Semesters verändert.

Durch den Vergleich beider Messungen wird eine mögliche Einstellungsänderung herausgestellt und mit einem zweiseitigen T-Tests auf Signifikanz überprüft.

Welch Two Sample t-test

```
data: rose1 and rose2
t = 0.3823, df = 36.847, p-value = 0.7044
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.8601482 1.2601482
sample estimates:
mean of x mean of y
1.8      1.6
```



Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

Die dem T-Test zugrundeliegende Hypothese besagt, dass die Mittelwerte der gemessenen Einstellungen der Studierenden zu Semesterbeginn und Semesterende gleich sind. Diese kann nicht abgelehnt werden, da die Überschreitungswahrscheinlichkeit  $p=0.7044$  beträgt. D.h. es findet keine signifikante Veränderung der Einstellungen der Studierenden in Bezug auf die Lernsoftware *Statistik interaktiv komplett* zu einem Signifikanzniveau von 5% oder 10% statt.

Zusätzlich wird auf der Laborseite **Hypothese15a** gemessen, ob sich die Fähigkeit der Studierenden, eine Statistiklaborseite eigenständig zu nutzen, signifikant verändert hat. Dies erfolgt im R-Kalkulator mit Hilfe eines zweiseitigen T-Tests. Der Output des R-Kalkulators, in dem die notwendigen Berechnungen vorgenommen werden, lautet wie folgt:

Welch Two Sample t-test

```
data: nutzung1 and nutzung2
t = 0.2708, df = 36.003, p-value = 0.788
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.3243949  0.4243949
sample estimates:
mean of x mean of y
 -0.95   -1.00
```

Die dem T-Test zugrundeliegende Hypothese besagt, dass die Mittelwerte der Antworten der Studierenden zu Semesterbeginn und Semesterende gleich sind. Diese kann nicht abgelehnt werden, da die Überschreitungswahrscheinlichkeit  $p=0.7044$  beträgt. D.h. es findet keine signifikante Veränderung der Bewertung der Fähigkeit der eigenständigen Nutzung der Studierenden in Bezug auf das Statistiklabor zu einem Signifikanzniveau von 5% oder 10% statt. Dies kann auch deskriptiv anhand der entsprechenden Stabdiagramme für Semesterbeginn und Semesterende auf der Laborseite ... betrachtet werden.

Außerdem werden auf den Laborseiten ... deskriptiv die folgenden Aspekte untersucht:

- eigenständige Nutzung des Statistiklabors nach gemeinsamer Erstellung von Statistiklaborseiten
- Selbstständigkeit im Statistiklabor (**Hypothese15a**)
- Erleichterung des statistischen Arbeitens durch das Statistiklabor (**Hypothese15b**)
- umfassende statistische Lösungsmöglichkeiten durch das Statistiklabor

Betrachten wir nun mögliche Zusammenhänge der drei dargestellten Merkmale in einer dreidimensionalen Kontingenztabelle (Laborseite: **Hypothese15c**):

	Erleichterung	1	1	
	Möglichkeiten	0	1	
Selbstständig				
0		0	0.05	0.05
1		0.05	0.90	0.95
		0.05	0.95	1.00

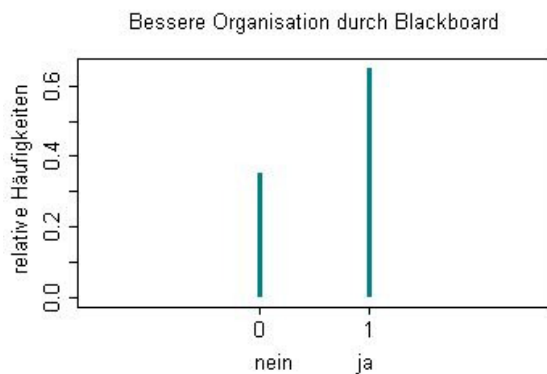
Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

Exemplarische Interpretation: 90% der Studierenden haben schon selbstständig Laborseiten erstellt, sind der Meinung, dass ihnen das Statistiklabor das statistische Arbeiten erleichtert und haben aufgrund ihrer bisherigen Kenntnisse den Eindruck, dass ihnen das Statistiklabor umfassende statistische Lösungsmöglichkeiten bietet.

Es fällt auf, dass in diesem Bereich keine signifikanten Veränderungen im Verlauf des Semesters gemessen werden konnten. Dies kann auch durch Schwierigkeiten der Studierenden bei der Beantwortung der Fragen zustande gekommen sein. Aus diesem Grund wird dieser Bereich stärker im Bereich der qualitativen Evaluation untersucht werden.

### 3.e) Evaluation des eingesetzten Lernmanagementsystems Blackboard

Abschließend wird der Einsatz der Lernsoftware im Rahmen des Lernmanagementsystems Blackboard evaluiert. Es wird untersucht, ob die Organisation des Lernens und der Umgang mit Neuen Medien durch den Einsatz in einem Lernmanagementsystem verbessert wird. Die Einschätzung der Studierenden zu Semesterbeginn ist auf der Laborseite **Hypothese16** in dem folgenden Stabdiagramm dargestellt:



Ein Großteil der Studierenden vertritt die Meinung, dass ihnen das Lernmanagementsystem eine bessere Organisation des Lernens ermöglicht. An dieser Stelle sollte darauf hingewiesen werden, dass die qualitative Evaluation ergeben hat, dass die Studierenden größtenteils nicht zwischen dem Lernmanagementsystem und den in diesem bereitgestellten multimedialen Lernmodulen differenzieren, was Auswirkungen auf die quantitativen Evaluationsergebnisse gehabt haben kann. Hier muss die qualitative Evaluation einsetzen.

In einem nächsten Schritt wird das Lernmanagementsystem hinsichtlich seiner Eignung zur Kombination traditioneller und multimedialer Veranstaltungskomponenten überprüft. Die Studierenden wurden zu Semesterbeginn und Semesterende um eine diesbezügliche Einschätzung gebeten. Eine deskriptive Analyse erfolgt auf der Laborseite **Hypothese17**. Außerdem wird das Ergebnis mittels eines t-Tests überprüft.

In einem letzten Schritt wird überprüft, ob das Lernmanagementsystem die Teamarbeit der Studierenden, sowie die Kommunikation zwischen Studierenden und Dozenten positiv unterstützt. Die diesbezüglich existierende Bewertung der Studierenden wird deskriptiv untersucht und grafisch in Form von Stabdiagrammen dargestellt.

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

Dabei werden die Bewertungen zu Semesterbeginn und –ende gegenübergestellt und verglichen (vgl. Laborseite **Hypothese18b**). Zusammenhänge der Merkmale werden mit Hilfe der Regressionsanalyse untersucht. Die Kommunikationsmöglichkeiten im Lernmanagementsystem werden sowohl zu Beginn des Semesters als auch zu Ende des Semesters als praktisch eingeschätzt. Der Vergleich der Stabdiagramme zeigt in der konkreten Anwendung keine nennenswerte Veränderung. Dieses Evaluationsergebnis wird durch den zweiseitigen T-Test bestätigt (vgl. Laborseite **Hypothese18**):

Welch Two Sample t-test

```
data: Frabo1 and Frabo2
t = -0.9883, df = 35.888, p-value = 0.3296
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.9558712 0.3295554
sample estimates:
mean of x mean of y
-1.263158 -0.950000
```

Das Lernmanagementsystem Blackboard wird im Vergleich zu einer „klassischen“ Internetkommunikation über Webseiten und E-Mail als besser eingeschätzt, um die notwendigen Maßnahmen durchzuführen. Dies lässt sich aus dem unteren Stabdiagramm erkennen.

#### 4. Qualitative Evaluation – Zusammenfassung der Befragung

Zu Semesterende wurde eine qualitative Befragung der Studierenden durchgeführt. Die Ergebnisse werden im Folgenden zusammengefasst wiedergegeben. Umfangreichere Ausführungen zu der Befragung kann der Leser dem Anhang dieser Arbeit entnehmen.

***Frage 1: Konnte Ihr Interesse für Statistik durch das Projekt geweckt werden? Wenn ja, ist dies eher auf den Einsatz von Multimedia zurückzuführen, oder auf andere Aspekte wie zum Beispiel die Betreuung, Form der Vorlesung, ...***

Die Antworten der Studierenden zeigten, dass bei einem Großteil das Interesse für Statistik geweckt werden konnte. Dies ist gleichermaßen auf drei Aspekte und deren Kombination zurückzuführen:

- die Gestaltung der Veranstaltung: Vorlesung, Übung,
- die Rechnerstunde in kleiner Gruppe und die Arbeit in Teams
- den Einsatz der multimedialen Lernsoftware Statistik interaktiv komplett, insbesondere des Statistiklabors,
- die starke Betreuung der Studierenden in Verbindung mit der Nähe zu den jungen Dozenten

Das Leitfadenprinzip wurde in diesem Zusammenhang gelobt, da es eine Organisation der verwendeten Module bereits vornahm. Viele betonten, dass besonders das Arbeiten im Labor, in der Vorlesung und zu Hause, ihr Interesse für die Statistik geweckt hat, sie das Statistiklabor als ein Instrument zur Arbeitserleichterung empfunden haben. Zusammenfassend musste festgestellt werden, dass die multimedialen Lernmodule zwar einen positiven Effekt auf das Interesse der Studierenden für die Statistik hatten, dass aber die starke Betreuung durch junge Dozenten deutlich höher bewertet wurde.

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität  
Bielefeld

**Frage 2: Empfanden Sie es als positiv, Teile der Prüfungsleistung schon während des Semesters erbringen zu können (Aufgaben, Reports) - wenn ja, warum? ... oder empfanden Sie die höhere Belastung während des Semesters als eher störend (Gründe)?**

Ein Großteil der Studierenden hat eine positive Bewertung vorgenommen. Es wurde betont, dass die Reports während des Semesters eine kontinuierliche Auseinandersetzung mit dem Stoff hervorgerufen haben, dies aber auch sehr arbeitsaufwendig war. Die Abgabe der wöchentlichen Übungsaufgaben wurde in diesem Zusammenhang ähnlich bewertet: die Belastung war groß, aber man setzte sich kontinuierlich mit dem Stoff auseinander. Interessant ist die Beobachtung, dass sich die ausländischen Teilnehmer durchweg positiv äußerten und nicht über eine zu hohe Belastung klagten, sondern das kontinuierliche Lernen hervorhoben. Einige Teilnehmer schlugen vor, den Abschlussreport einige Wochen früher, also vor Semesterende zu schreiben, um ihn aus der Vorbereitung für die anderen Klausuren rauszuhalten.

**Frage 3: Das Projekt besteht aus verschiedenen Komponenten: Vorlesung, Übung, Rechnerstunden, selbstständiges Reportwriting – welche Meinung haben Sie bezüglich der einzelnen Komponenten (Stichworte), wo haben Sie am meisten gelernt? Verbesserungsvorschläge/Kritik?**

Einige Studierende bewerteten alle Komponenten in Verbindung miteinander und betonten, dass das Prinzip Vorlesung-Übung-Rechnerstunde-Reportwriting sehr hohe Lernerfolge hervorgerufen hat, das Zusammenwirken und die zeitliche Abfolge (langsame Vertiefung, schrittweises selbständiges Arbeiten) der einzelnen Komponenten. Viele betonten, dass besonders die Unterstützung in der Rechnerstunde dazu beigetragen hat, dass sie den Umgang mit dem Statistiklabor erlernten und die zur Lösung der statistischen Probleme einsetzen konnten. Zusammenfassend kann festgestellt werden, dass die Rollen der einzelnen Komponenten erkannt und positiv bewertet wurden. Wünschenswert hätten es einige gefunden, ein größeres Angebot von betreuten Rechnerstunden angeboten bekommen zu hätten.

**Frage 4: Würden Sie sich wiederholt für ein Multimediaprojekt anmelden, wenn Sie die Möglichkeit dazu hätten? (Begründung)**

Den Antworten war zu entnehmen, dass ein Großteil der Studierenden jederzeit wieder an einem Multimediaprojekt teilnehmen würde. Es wurde häufig nach einem Anschlusskurs Statistik interaktiv II gefragt. Eine häufig genannte Begründung ist das interessantere Lernen durch eigene Aktivität in den Komponenten der Lernmodule. Ein weiterer Grund ist die Arbeit in kleinen Gruppen, die von den Studierenden in Verbindung gebracht wird mit einem Multimediaprojekt. „Angenehmes Lernen durch gute Betreuung, man fühlt sich nicht so sehr allein mit Problemen wie in vollen Vorlesungen.“ „Ja, weil es schön war, auch mal in einer kleinen Gruppe zu lernen und weil man viel mehr Kontakt zum Dozenten hatte.“ Einige betonten, dass sie großes Interesse an Multimediaprojekten haben, da sie der Meinung sind, besser für die berufliche Zukunft und den damit verbundenen Umgang mit neuen Medien vorbereitet zu werden. Ein interessanter Beitrag lautet wie folgt: „Ja, da der Umgang mit Medien immer wichtiger wird; Bearbeitung ist viel einfacher (Grafiken etc.) und spart viel Zeit, die für wichtigere Erklärungen verwendet werden können.“ Dieses Ziel wurde mit dem Einsatzkonzept verfolgt, es scheint sich erkennbar durchgesetzt zu haben. Die Prüfung in Form der statistischen Reports wird von vielen als weiterer Grund genannt, sich wieder für ein solches Projekt anzumelden.

**Frage 5: Hat Ihnen das Reportwriting das wissenschaftliche Arbeiten nähergebracht?**

Die Antworten lassen darauf schließen, dass die Studierenden mit dem Begriff „wissenschaftlichem Arbeiten“ Schwierigkeiten hatten. Einige betonten, dass sie gelernt hätten, ihre Lösungen zu strukturieren, Methoden ausgewählt und begründet anzuwenden. Außerdem bemerkten einige, dass sie gelernt hätten, ein Problem zu beschreiben, einzugrenzen und Analysen in Bezug auf das Problem zu interpretieren.

**Frage 6: Beurteilen Sie die einzelnen Komponenten der eingesetzten Multimediastftware.**

Die Beurteilung erfolgte komponentenweise, wie die folgende Dokumentation der Antworten zeigt:

- Glossar: technische Probleme
- Beispiele: gut für Anwendungsverständnis
- Applets: gute Anschauungskraft
- Statistiklabor: es werden technische Mängel angesprochen,

aber ansonsten wird das Statistiklabor als eine sehr wichtige Komponente des statistischen Problemlösens betrachtet. Ein Studierender macht Vorschlag, eine Formelsammlung für den R-Kalkulator zu entwickeln und zu implementieren. Obwohl an dieser Stelle nach den Multimediakomponenten gefragt wurde, bemerkten viele Studierende, dass ihnen die Vorlesungen und Übungen am meisten gebracht hätten. Der Umgang mit dem R-Kalkulator wurde von vielen als schwierig bezeichnet, sie betonten, mehr Hilfestellungen zu benötigen, um mit dem R-Code hantieren zu können. Die Studierenden erkannten, dass sich das Labor noch in keiner Endversion befand und betonten, dass es das wichtigste Element der Lernsoftware sein wird, sobald alle technischen Mängel

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität Bielefeld

beseitigt wären. Die allgemeinen Bemerkungen der Studierenden waren nicht einheitlich: einige bezeichneten die Lernmodule als übersichtlich, hoben die Vernetzung der Komponenten durch Hyperlinks positiv hervor, andere kritisierten gerade diese Aspekte, schlugen vor, sich auf das Wesentliche zu beschränken, sprachen sich gegen die Hyperlinks aus.

**Frage 7: Welche der in der Vorlesung vermittelten statistischen Gebiete finden Sie am interessantesten z. B. weil Sie sich vorstellen können, sie später noch gebrauchen zu können?**

Die Antworten waren sehr unterschiedlich: ein Teil der Studierenden erachtete alle Methoden in Verbindung miteinander als notwendig, merkte z.B. an, dass die Normalverteilung allein nicht viel Erkenntnis bringt, wenn man sich nicht mit Hilfe deskriptiver Methoden einen Überblick über die Daten verschaffen kann. Andere griffen einzelne Gebiete heraus und begründeten, warum sie gerade diese als interessant und später relevant (Berufsleben) erachteten.

**Frage 8: Wie schätzen sie das Teamwork ein? (Begründung)**

Ein Großteil der Antworten spiegelte eine sehr positive Bewertung des Teamworks wieder. Der Aspekt des gegenseitigen Ergänzens, insbesondere in Bezug auf die Erstellung der ersten beiden Reports tauchte in vielen Antworten auf. „Gut, es ist auch für später wichtig, im Team arbeiten zu können. Man kann sich gegenseitig helfen und Probleme diskutieren. Die Arbeitsergebnisse sind besser.“ Die Studierenden berichteten über das gegenseitige Aufdecken und Beseitigen von Fehlern und Verständnisschwierigkeiten, Sammeln von Ideen zur geeigneten Lösung statistischer Problemstellungen. Interessant sind auch Bemerkungen dahingehend, dass nicht nur innerhalb der gebildeten Zweierteam, sondern auch zwischen verschiedenen Teams Kommunikation stattfand.

**Frage 9: Sehen Sie die Erstellung der statistischen Reports als gute Einführung in das statistische Problemlösen? (Begründung)**

Die Antworten der Studierenden waren durchgehend positiv. Sie bemerkten, dass sie durch die Anfertigung der Reports gelernt hätten, wie sie an das Lösen (statistischer Probleme) herangehen sollten. „Ja, man befasst sich mit einem Problem, entscheidet über Methoden, dieses zu lösen und kommt schließlich zu einem Lösungsvorschlag.“ „Ja, denn es ist nicht einfach, Daten systematisch zu analysieren und erfordert viel Übung.“ „Ja, da man sich so umfangreich Gedanken um ein Problem macht und dieses auch für den Alltag sehr interessant sein kann.“ Außerdem erlernten sie, strukturierte Lösungen zu erstellen, Analysen zu interpretieren und ihr Arbeiten zu dokumentieren.

## 5. Rückbezug auf die Realität

In einem statistischen Report werden die erzielten Analyseergebnisse abschließend im Sinne der Problemstellung interpretiert. Ein solcher Rückbezug zur Realität erscheint hier nur bei der qualitativen Evaluation sinnvoll, da die Ergebnisse der quantitativen Evaluation auf Basis eines geringen Stichprobenumfangs erzielt wurden. Die schriftliche Befragung wurde dem Leser bereits überblickartig vorgestellt. Das übergeordnete Ziel der Evaluation besteht darin, Schwächen und Stärken der Lernsoftware und des Einsatzkonzeptes zu identifizieren und eine schrittweise Optimierung beider anzustreben. Aus den Antworten der Studierenden können folgende **Evaluationsergebnisse** abgeleitet werden:

- Das Interesse für Statistik konnte durch den Einsatz der Lernsoftware in Verbindung mit der starken Betreuung der Studierenden geweckt werden. Viele Studierende betonten, dass sie durch den Einsatz des multimedialen Lernmodule motivierter an die statistischen Inhalte herangegangen sind, insbesondere bei der Nacharbeitung der Vorlesungen.
- Das im Einsatzkonzept verankerte Leitfadenprinzip scheint ein geeignetes Instrument zum Guiding der Studierenden durch die Lernmaterialien darzustellen. Dem Statistiklabor kommt eine zentrale Rolle zu. Die Befragung hat verdeutlicht, dass die Studierenden das Statistiklabor als ein Instrument zur Arbeitserleichterung empfunden haben.

Forschungsprojekt NEUE STATISTIK – Abschlussreport der Evaluation im SoSe 2003 der Universität Bielefeld

- Allerdings ist aus den Antworten deutlich herauszuhören, dass der Hauptgrund für ihr Interesse an der Statistik in der Betreuung liegt: „Die Atmosphäre in den Vorlesungen war sehr angenehm und die Dozenten haben super erklärt.“ „... lockerte Form der Vorlesung, Arbeiten im Team“. Die angenehme Atmosphäre wird insofern gelobt, als dass man sich nicht scheute, Fragen zu stellen. Die kleine Gruppe, die eine starke Betreuung des Einzelnen ermöglichte, wurde positiv hervorgehoben. Durch die im Einsatzkonzept verankerten Aktivitäten der Studierenden während des Semesters wurde ein kontinuierliches Lernen hervorgerufen und unterstützt. Die Abgabe der wöchentlichen Übungsaufgaben wurde in diesem Zusammenhang ähnlich bewertet: die Belastung war groß, aber man setzte sich kontinuierlich mit dem Stoff auseinander. Kritisch sollte allerdings die Kritik einiger Teilnehmer betrachtet werden, die bemerkten, dass durch den hohen Arbeitsaufwand in Statistik andere Fächer zu kurz gekommen sind. Einige Teilnehmer schlugen vor, den Abschlussreport einige Wochen früher, also vor Semesterende zu schreiben, um ihn aus der Vorbereitung für die anderen Klausuren rauszuhalten. Ein weiterer Vorschlag bestand darin, die wöchentlich abzugebenden Lösungen der Übungsaufgaben stärker gewichtet in die Endnote einzubeziehen, da diese einen großen Teil der Arbeit ausmachten.
- Viele betonten, dass besonders die Unterstützung in der Rechnerstunde dazu beigetragen hat, dass sie den Umgang mit dem Statistiklabor erlernten und die zur Lösung der statistischen Probleme einsetzen konnten. Positiv wurde außerdem hervorgehoben, dass die Aufgaben in der Vorlesung und Übung gemeinsam mit dem Dozenten erarbeitet wurden, also im Dialog gelernt wurde. Einige Studierende hätten in der Übung gern mehr im Statistiklabor gearbeitet - es bleibt zu überlegen, auch die Übung in einem Rechnerpoolraum abzuhalten, um Aktivität der Studierenden am Rechner zu ermöglichen. Kritisch wurde die Fülle des zu vermittelnden und zu erlernenden Stoffes angesprochen. Die Studierenden fühlten sich durch die wöchentlich neu vermittelten statistischen Inhalte/Methoden überfordert.

Das Evaluationsinstrument wurde im Einsatz einem Pretest unterzogen. Der resultierende Evaluationsreport wurde bereitgestellt. In diesem befanden sich nur exemplarisch konkrete Interpretationen zur Evaluationsstudie. Dies ist auf den geringen Stichprobenumfang und damit die fehlende Repräsentativität der erzielten Ergebnisse zurückzuführen. Die Evaluationsstudie fand auf Basis einen geringen Stichprobenumfangs statt. Dies bot den Vorteil, eine bessere Kommunikation mit den Teilnehmern zu erreichen und war zudem durch die technischen und organisatorischen Rahmenbedingungen vorgegeben.

Eine weitere Möglichkeit hätte darin bestanden, den konkreten Evaluationsreport als Beispiel bereitzustellen. Dies ist nicht geschehen, er wurde in dem beschriebenen Entwicklungsprozess lediglich als Methode oder Werkzeug eingesetzt. Das Werkzeug ermöglichte die kritische Betrachtung der Arbeitshypothesen und die daraus resultierende Eliminierung oder ihre mögliche Neugenerierung.

## 14 Ergebnisse und bleibende Erkenntnisse

Die gesammelten Erfahrungen und Ergebnisse der beiden Durchführungen der Evaluationsstudie werden zusammengefasst dargestellt. Parallel werden bleibende Erkenntnisse für den zukünftigen Einsatz multimedialer Lernsoftware in der Statistikausbildung abgeleitet. Die Ausführungen werden nach zwei Themengebieten gegliedert dargestellt:

- Gestaltung der Materialien
- Entwicklung des Einsatzkonzeptes

An dieser Stelle werden nur die zentralen Erkenntnisse dargestellt. Die ausführliche Darstellung der Evaluationserfahrungen hat in den entsprechenden Kapiteln zu den Durchführungen der Evaluationsstudie stattgefunden.

### 14.1 Gestaltung der Materialien

Die Gestaltung der Materialien wurde entscheidend durch die Überlegenheit des *guided-learning*s gegenüber dem *self-paced-learning* beeinflusst. Besonders in Bezug auf die Vermittlung des Verfassens statistischer Reports wurden Erfolge mit zusätzlich entwickelten Materialien (Musterreports, Animationen, Coachingkomponente) erzielt. Die Methode des *guided-learning*s sollte zukünftig sowohl bei der Entwicklung von Lernsoftware als auch bei der Entwicklung von Einsatzkonzepten im Vordergrund stehen.

**Zusammenfassende Bewertung ausgewählter Komponenten** Das Statistiklabor war die zentrale Komponente der Lernsoftware. Sie führte den Studierenden die Bedeutung von Computational Statistics vor Augen, ermöglichte ihnen erste statistische Analysen, bildete die Basis für das Verfassen statistischer Reports. Die im Statistiklabor implementierte Berichterstellungsfunktion ermöglichte die Kombination von statistischen Analyseergebnisse und zugehöriger Interpretation und Dokumentation. Beide Durchführungen kamen zu einer positiven Bewertung des Statistiklabors hinsichtlich der folgenden Aspekte:

- Unterstützung beim statistischen Problemlösen
- erster Umgang mit Programmieren
- Erstellung statistischer Reports mit Hilfe der Berichterstellungsfunktion

Die Idee, das Statistiklabor zunächst zur Aufgabenlösung, dann schrittweise zur Erstellung statistischer Reports einzusetzen, hat sich als positiv erwiesen: Das statistische Problemlösen wurde etabliert und vermittelt, große Lernerfolge innerhalb des jeweiligen Semesters waren erkennbar. Hier war ein hoher Arbeitsaufwand erforderlich, was sich sowohl auf die Betreuung als auch auf die Korrekturen bezieht. Die Integration der Coachingkomponente *Statistisches Reportverfassen* sowie des Musterreports in die Lernsoftware werden den Betreuungsaufwand in Zukunft nur bedingt verringern können.

Zusammenfassend kann die gemeinsame Wirkung der Komponenten Vorlesung, Übung, Rechnerstunde, Statistiklabor, Reports als positiv bezeichnet werden. Die im Einsatzkonzept etablierte Kombination hat sich bewährt, ist daher für zukünftige Einsätze empfehlenswert. Es ist allerdings notwendig, eine stabile Version des Statistiklabors zu entwickeln, bevor es zu einem flächendeckenden Einsatz der Lernsoftware *Statistik interaktiv komplett* kommt. Die technischen Probleme, mit denen die Studierenden bei der jetzigen Version konfrontiert wurden, behinderten den Lernprozess unnötig. Der Aufbau, das Layout sowie die bereitgestellten Funktionalitäten der Laborobjekte wurden allerdings durchgehend positiv bewertet und sollten beibehalten werden.

Neben dem Statistiklabor kam den Theoriekomponenten eine tragende Rolle zu. Die Theoriekomponenten wurden durchgehend positiv bewertet. Sie eignete sich zum Einsatz in der Vorlesung, als auch zur eigenständigen Vor-/Nachbereitung dieser durch die Studierenden. Die Theoriekomponenten sind in ihrer jetzigen Version ein Startpunkt zu allen anderen Komponenten der Lernsoftware. Es können Beispiele, Aufgaben, Applets, Animationen und Musterreports von den Theoriekomponenten aus über Verlinkungen erreicht werden. Der Einsatz hat gezeigt, dass sich dieser Aufbau eignete, um die Studierenden zur Nutzung der einzelnen Komponenten zu motivieren. Es bleibt kritisch zu bemerken, dass die Theoriekomponenten nicht problemlos austauschbar sind, da die Beispiele stark integriert sind und keine unabhängigen Motivationsteile vorhanden sind. Bei der zukünftigen Entwicklung von Lernsoftware sollte auf stärkere Modifizierbarkeit der Komponenten geachtet werden.

Betrachtet man die Erfahrungen bezüglich der Komponente *Fallstudie*, so sollte über eine andere Form des Einsatzes nachgedacht werden. Die Komponente wurde nur dann genutzt, wenn sie in der Vorlesung eingesetzt wurde. Es bestand kein Interesse seitens der Studierenden, sich neben der Vorlesung mit den Fallstudien auseinanderzusetzen. Dies kann mehrere Gründe haben. Zum einen kann es an der Gestaltung der Fallstudien oder ihrem Inhalt liegen. Zum anderen kann es darauf zurückgeführt werden, dass die anderen Komponenten des Einsatzkonzeptes bzw. der Lernsoftware die Bedürfnisse der Studierenden bereits abgedeckt haben. Die Vorlesung könnte ein Grund für die geringe Nutzung der Fallstudien durch die Studierenden sein, d.h. in der Vorlesung wurden die Studierenden bereits für die statistischen Inhalte motiviert. Diese Aufgabe sollten eigentlich die Fallstudien übernehmen.

Die Applets und Animationen wurden verschieden stark eingesetzt. In der Vorlesung konnte gut auf Basis der Applets vermittelt werden, während die Animation zur Nacharbeitung besser geeignet war.

**Evaluationsinstrument** Wird das entwickelte Evaluationsinstrument flächendeckend eingesetzt, bietet es sich an, die Evaluationsergebnisse durch die Bildung von Kontrollgruppen (traditionell unterrichtete Gruppen) zu überprüfen. Dies ist jetzt nicht möglich, da die traditionellen Veranstaltungen völlig anders ausgerichtet sind. Man befindet sich auf diesem Gebiet noch in einer Evaluationsphase. Das Evaluationsinstrument stellt einen ersten Schritt dar, flächendeckend auf Basis großer Stichprobenumfänge evaluieren zu können. In diesem Zusammenhang sollte auch das Tutorial für das Statistiklabor schrittweise evaluiert werden.

## 14.2 Entwicklung des Einsatzkonzeptes

**Grad der Betreuung** Die Evaluationsstudie zeigte, dass die Lernerfolge der Studierenden insbesondere auf eigene Aktivitäten im Statistiklabor zurückgeführt werden können. Dies sollte zukünftig bei dem Einsatz von *Statistik interaktiv komplett* berücksichtigt werden. Die qualitative Befragung zu Semesterende zeigte, dass eine Durchführung der Übung im Rechnerpool von Vorteil wäre, um die Aktivität der Studierenden zu steigern.

**Tutorial** Es kann darüber nachgedacht werden, ob das Tutorial oder die Sammlung mehrerer Tutorials einen Teil einer Benutzereinführung darstellen sollten. Es besteht die Möglichkeit, weitere Materialien, z.B. dokumentierte Statistiklaborseiten einzusetzen, um dem Benutzer die Möglichkeiten des Statistiklabors zu demonstrieren.



**Konstruktivistisch geprägte Komponenten** Die 2. Durchführung der Evaluationsstudie hat gezeigt, dass Aktivität und Wissenskonstruktion seitens der Studierenden durch das Leitfadensprinzip hervorgerufen und gefördert werden. Auch die mündliche Prüfung hat gezeigt, dass sie eine geeignete Komponente darstellen kann, konstruktivistische Prozesse im Kleinen anzustoßen, alternative Lösungsstrategien aktiv zu erarbeiten. Die Methode des teilnehmenden Beobachtens zeigte, dass der Aktivitätsgrad der Studierenden bei der Bearbeitung der Übungsaufgaben und der Erstellung der statistischen Reports am höchsten war. Es erfolgte eine Diskussion der drei Reports in den Übungen. Diese hat sich als positiv erwiesen: Lösungsalternativen wurden diskutiert, Fehler aufgedeckt, Missverständnisse herausgestellt. Bei zukünftigen Einsätzen sollten Reportbesprechungen eine feste Komponente der Übungen darstellen.

**Lernmanagementsysteme** Der Einsatz eines Lernmanagementsystems hatte viele Vorteile, war allerdings auch mit erheblichem Aufwand verbunden. Der Dozent sollte vor der Entscheidung für ein Lernmanagementsystem die Informationsströme seiner Veranstaltung darstellen und überprüfen, ob sie den Einsatz eines Lernmanagementsystems verlangen, welche Anforderungen daran gestellt werden und ob es ihnen gerecht wird. Die Informationsströme können aus den Tätigkeiten der Beteiligten (Dozent, Tutor, Studierende/Teams) abgeleitet werden. Das der Veranstaltung zugrunde liegende Einsatzkonzept legte diese Tätigkeiten der einzelnen und somit die Informationsströme zwischen ihnen fest. Die durch das hier entwickelte Einsatzkonzept entstandenen Informationsströme sprachen für den Einsatz des Lernmanagementsystems Blackboard.<sup>317</sup>

**Auswahl, Training und Koordination des Personals** Es wurde deutlich, dass insbesondere der Einsatz des Statistiklabors mit einem hohen Betreuungsaufwand verbunden ist. Die Betreuung in Vorlesung, Übung, Rechnerstunde und Sprechstunde war für ein effizientes Arbeiten der Studierenden im Statistiklabor notwendig. Für zukünftige Einsätze, insbesondere bei einer höheren Teilnehmeranzahl, muss eine intensive Betreuung geplant werden. Die betreuten Rechnerstunden haben sich als wichtigste Betreuungskomponente herausgestellt.

---

<sup>317</sup>Vgl. 13.1.2.

## Fazit und Ausblick

Ausgangspunkt dieser Arbeit bildeten Defizite in der Statistikausbildung. Das Ziel bestand darin, diesen durch die neuen Medien positiv entgegenzuwirken. Dazu wurde die Statistik im ersten Teil charakterisiert. Es wurden historische und aktuelle Artikel herangezogen, um die Entwicklung und Schwierigkeiten der Statistikausbildung darzustellen. Zudem wurde gezeigt, dass der Konstruktivismus eine entscheidende Rolle in den Neugestaltungen der Statistikausbildung spielt. Es wurde eine Defizitklassifikation vorgenommen und Beziehungen zwischen den Defizitgruppen in der Statistikausbildung dargestellt. Dabei wurden drei Gruppen identifiziert:

- Übergeordnetes Defizit: Die Unterschätzung der Disziplin Statistik kann häufig darauf zurückgeführt werden, dass nicht erkannt wird, wie komplex sie ist, d.h. dass sie sich aus mehr als statistischen Methoden zusammensetzt.
- Ursachendefizite
- Folgedefizite

Die Defizitklassifikation bildete den Ansatzpunkt für Neugestaltungsmaßnahmen, die mit Hilfe der neuen Medien durchgeführt werden sollten. Anschließend wurde anhand exemplarischer Umsetzungen gezeigt, wie die multimediale Lernsoftware *Statistik interaktiv komplett* eingesetzt werden könnte, um einzelnen Defiziten positiv entgegenzuwirken. Dabei fand eine Integration konstruktivistischer Prinzipien statt. Die ausgearbeiteten Lösungsvorschläge hatten Demonstrationscharakter. Insbesondere die Einsatzmöglichkeiten und Funktionalitäten des Statistiklabors wurden herausgestellt.

Im zweiten Teil wurden Definitionsversuche zu den Begriffen *Multimedia* und *Neue Medien* vorgestellt. In Abgrenzung hierzu wurde der Versuch einer für die Arbeit gültigen Begriffsdefinition erarbeitet. Diese stellt die Interaktion des Anwenders mit dem System in den Mittelpunkt. Der Anwender hat durch die neuen Medien viele Interaktionsmöglichkeiten und konstruiert sein Wissen durch Aktivität.

Facetten der Diskussion um den Einsatz und die Evaluation multimedialer Lernsoftware wurden kritisch dargestellt. Für die Evaluationsstudie dieser Arbeit wurden folgende zentrale Erkenntnisse und Konsequenzen abgeleitet.

- Der gezielte Einsatz und die Evaluation multimedialer Lernsoftware müssen organisiert werden. Es bedarf der Entwicklung eines Einsatzkonzeptes für *Statistik interaktiv komplett*.
- Es wurde die Entscheidung begründet, in der Evaluationsstudie phasenweise zu evaluieren. Die Evaluation soll dabei als Bewertung verstanden werden. Es wurde die Konsequenz abgeleitet, bei der Bewertung qualitative und quantitative Evaluationsverfahren zu kombinieren.
- Der Entwicklungsprozess multimedialer Lernsoftware spielt bei der Bekämpfung der Defizite in der Statistikausbildung eine entscheidende Rolle. Entwicklung, Einsatz und Evaluation multimedialer Lernsoftware müssen parallel betrachtet werden.
- Eine Evaluation mittels Kriterienkatalogen wird ausgeschlossen. Die Evaluationsstudie muss das Ziel verfolgen, individuelle Kriterien für konzipierte Einsätze abzuleiten und hinsichtlich ihrer Eignung empirisch zu untersuchen. Die beschriebenen Zielkompetenzen der Studierenden bildeten dabei den zu erreichenden Sollzustand.

Als Konsequenz wurde im dritten Teil untersucht, welche Aspekte bei der Entwicklung und Gestaltung multimedialer Lernsoftware zu beachten sind. Zunächst wurde die Methode des Prototypings in das Gebiet "Softwareengineering" eingeordnet. Diese wurde auf die Entwicklung der Lernsoftware *Statistik interaktiv komplett* angewendet. Es wurde herausgearbeitet, dass ein dynamischer Entwicklungsansatz benötigt wird. Auf Basis des evolutionären Gestaltungsansatzes des Softwareengineering wurden ein Vorgehensmodell und ein Architekturmodell zum Prototyping multimedialer Lernsoftware entwickelt. Das Vorgehensmodell sieht eine phasenweise Entwicklung der Lernsoftware und eines Einsatzkonzeptes für diese vor. Am Beispiel *Statistik interaktiv komplett* wurden anschließend einzelne Entwicklungsphasen vorgestellt. Ziele der Entwicklung waren:

- Zerlegung der Statistik in einzelne Module, die zusammengesehen ein einheitliches Bild der Statistik darstellen, d.h. eine modulare Gestaltung der Lernsoftware
- Vermittlung der Statistik als eine Disziplin, die sich aus statistischem Problemlösen, statistischem Denken und statistischen Methoden zusammensetzt
- Gestaltung der Module/Komponenten und des Einsatzkonzeptes nach konstruktivistischen Prinzipien
- Sicherstellung der Prinzipien Corporate Identity, Nachhaltigkeit und Modifizierbarkeit der Lernsoftware, um dem flächendeckenden Einsatz gerecht zu werden

Anhand der Komponente *Statistiklabor* wurde der phasenweise Entwicklungsprozess<sup>318</sup> exemplarisch demonstriert. Es wurden folgende Meilensteine bezüglich der Weiterentwicklung der Laborobjekte und -funktionalitäten herausgestellt:

- Ausgangspunkt  
Übernahme des Labors aus *Statistik interaktiv*. Das Labor der Vorgängersoftware wurde zunächst übernommen. Diese hatte allerdings erhebliche Mängel, z.B. keine stabile Lauffähigkeit, die schrittweise verbessert wurden.
- Entwicklung einer benutzerfreundlichen Oberfläche  
Die Benutzeroberfläche wurde schrittweise an die den Benutzern bekannten Windowsstandards angeglichen.
- Bereitstellung und Dokumentation von Laborseiten  
Um Transparenz und Nachvollziehbarkeit zu gewährleisten, empfiehlt sich eine Dokumentation jeder im System bereitgestellten Laborseite.
- Entwicklung und Implementierung der Berichterstellungsfunktion  
Die Berichterstellungsfunktion ermöglicht die Kombination statistischer Analysen und ihrer Interpretationen zu einem statistischen Report.
- Entwicklung fehlerfrei funktionierender Laborobjekte

Ein weiterer Schwerpunkt wurde auf die Beurteilung der entwickelten Lernmodule gelegt. Diese wurde durch den Einsatz der Module durchgeführt. Außerdem wurde der Konstruktivismus als Messlatte herangezogen. Anhand von Beispielen wurde überprüft, inwiefern konstruktivistische Ideen in der Lernsoftware *Statistik interaktiv komplett* und in ihrem parallel entwickelten Einsatzkonzept umgesetzt wurden. In einer kritischen Reflektion wurde abschließend dargestellt, dass diese Ziele einen erwünschten Idealzustand der Entwicklung darstellten und nur bedingt umgesetzt werden konnten. Ein gravierender Kritikpunkt war der folgende:

Die Lernmodule von *Statistik interaktiv komplett* können nicht modifiziert werden. Es ist dem Dozenten z.B. nicht möglich, Theoriekomponenten durch individuelle Teile zu ergänzen. Die Modifizierbarkeit ist lediglich im Statistiklabor gewährleistet. Allerdings ist die Bereitstellung vorlesungsspezifischer Laborseiten nur durch die Nutzung eines Lernmanagementsystems oder zusätzlichen Laufwerkes möglich. Der Einsatz der Lernsoftware erfolgt im Rahmen eines Lernmanagementsystems.

Der abschließende Teil berichtete reflektierend über die beiden Durchführungen der Evaluationsstudie, in denen sowohl die Komponenten der Lernsoftware *Statistik interaktiv komplett* als auch das für sie entwickelte Einsatzkonzept in unterschiedlichen Versionen empirisch überprüft wurden. Zunächst erfolgten allgemeine Ausführungen zu Einsatz- und Evaluationskonzepten. Es wurde gezeigt, dass diese unter Berücksichtigung verschiedener Einsatzszenarien entwickelt werden müssen. Anschließend fand die Abgrenzung zwischen qualitativer und quantitativer Evaluation für die vorliegende Arbeit statt. Gemeinsamkeiten der Einsatzkonzepte Ia und Ib, die den beiden Durchführungen der Evaluationsstudie zugrunde liegen, wurden herausgestellt.

<sup>318</sup>Die Projektgruppe Bielefeld war eine der Projektgruppen, die zusammen mit CeDiS das Statistiklabor weiterentwickelten.

In der ersten Durchführung der Evaluationsstudie wurde ausschließlich qualitativ, in der zweiten vorwiegend quantitativ evaluiert. Die zweite Evaluationsstudie integrierte qualitative Elemente dort, wo die quantitative Evaluation an Grenzen stieß. Es konnten die folgenden Erkenntnisse aus der Evaluationsstudie abgeleitet werden:

- Das Verfassen statistischer Reports stellt eine geeignete Möglichkeit dar, statistisches Problemlösen zu erlernen und zu dokumentieren. Es kombiniert statistische Analysen und die Interpretation erzielter Ergebnisse. Durch die Reporterstellung wird transportiert, dass der Rechner ein zielführendes Instrument der statistischen Problemlösung darstellen kann. Der statistische Report eignet sich sowohl als Vermittlungs- als auch Bewertungsinstrument. Es empfiehlt sich, Prüfungsordnungen so umzugestalten, dass Bewertungsformen dieser Art eingesetzt werden können.
- Die Methode des *guided-learnings* sollte zukünftig sowohl bei der Entwicklung von Lernsoftware als auch bei der Entwicklung von Einsatzkonzepten im Vordergrund stehen.
- Für zukünftige Einsätze muss eine intensive Betreuung der Teilnehmer eingeplant werden. Die betreuten Rechnerstunden haben sich als wichtigste Betreuungskomponente herausgestellt. Es wurde deutlich, dass die kleine Gruppe und die damit verbundene höhere Betreuung des einzelnen Studierenden zu Lernerfolgen führte.
- Der Einsatz eines Lernmanagementsystems hatte viele Vorteile, war allerdings mit erheblichem Aufwand verbunden. Der Dozent sollte vor der Entscheidung für ein Lernmanagementsystem die Informationsströme seiner Veranstaltung darstellen. Er muss überprüfen, ob sie den Einsatz eines Lernmanagementsystems verlangen, welche Anforderungen an ein solches System gestellt werden und ob es den Anforderungen gerecht wird.
- Die Organisation und Durchführung der multimedialen Statistikveranstaltung in einem Lernmanagementsystem führte zu einem neuen Defizit. Die Studierenden müssen den Umgang mit diesem sowie der Lernsoftware erlernen, bevor sie sich dem Inhalt Statistik widmen.

Während der 2. Durchführung der Evaluationsstudie fand die Entwicklung eines Evaluationsinstrumentes statt. Mit dessen Hilfe können die multimediale Lernsoftware sowie ihr Einsatz evaluiert werden. Das Evaluationsinstrument bietet den Einsatz statistischer Analyseverfahren an, konzentriert sich auf die quantitative Evaluation. Die Umsetzung der Analyseverfahren fand auf speziell entwickelten Statistiklaborseiten statt. Die Defizite und daraus abgeleiteten Zielkompetenzen des Studierenden führten zu den folgenden Untersuchungsschwerpunkten:

1. statistisches Reportverfassen, statistisches Verständnis, Methodenverständnis
2. Adaptivität multimedialer Komponenten
3. Gestaltung des Einsatzkonzeptes

Auf Basis der Zielkompetenzen und Untersuchungsschwerpunkte erfolgte die Formulierung von Arbeitshypothesen. Jede Arbeitshypothese wurde entweder qualitativ oder quantitativ untersucht. Es erfolgte die Bereitstellung eines Evaluationsreports. Dieser bestand aus den Laborgrafiken, sowie Interpretations- und Auswertungsmodulen inferenzstatistischer Methoden. Um das Verfassen eines Evaluationsreports zu demonstrieren, wurden exemplarisch konkrete Interpretationen vorgenommen.

**Ausblick** Wird das Evaluationsinstrument flächendeckend eingesetzt, bietet es sich an, die Evaluationsergebnisse durch die Bildung von Kontrollgruppen<sup>319</sup> zu überprüfen. Dies war bis jetzt nicht möglich, da die traditionellen Veranstaltungen anders ausgerichtet waren als die multimedialen. Man befindet sich auf diesem Gebiet noch in einer Evaluationsphase. Das Evaluationsinstrument stellt einen ersten Schritt dar, flächendeckend auf Basis großer Stichprobenumfänge evaluieren zu können. Setzen viele Universitäten die Software auf Basis des entwickelten Einsatzkonzeptes gleichzeitig ein und evaluieren mit Hilfe des Evaluationsinstrumentes, können die erhobenen Evaluationsdaten zusammengefasst und evaluiert werden. Die Bereitstellung eines Evaluationsinstrumentes standardisiert somit den flächendeckenden Einsatz bzw. die Evaluation. Zukünftige Projekte sollten kritisch hinterfragen, was das Ziel flächendeckender Einsätze ist. Der Autor vertritt die Meinung, dass nur bestimmte Informationen durch einen flächendeckenden Einsatz bzw. eine Evaluation gewonnen werden können.

In diesem Zusammenhang sollte ein Tutorial für das Statistiklabor schrittweise entwickelt und evaluiert werden.

Im Rahmen der 2. Durchführung wurden Versuche unternommen, mit der Entwicklung eines Tutorial für das Statistiklabor zu beginnen. Das Ziel bestand darin, die Studierenden durch ein Tutorial in die Lage zu versetzen, statistische Probleme mit Hilfe des Statistiklabors zu lösen. Es sollte zusätzlich keine externe Betreuung notwendig sein. Es folgen idealtypische Überlegungen zur Entwicklung eines Tutorials.

Das Tutorial sollte abgegrenzt werden zu einer Hilfe, die alle Objekte des Statistiklabors und ihre Funktionalitäten beschreibt. Der Unterschied zwischen einer Hilfe und einem Tutorial besteht darin, dass das Tutorial zum statistischen Problemlösen anleitet, anstatt allein zur Benutzung der Software. Damit schrittweise vorgegangen werden kann, empfiehlt es sich, mehrere Tutorials (eine Sammlung von Tutorials) zu erstellen. Diese sollten Folgendes leisten:

- Vermittlung schrittweisen statistischen Arbeitens, Anleitung zur Verwendung des Statistiklabors und zum methodischen Vorgehen
- Demonstration der Kombinationen der Laborobjekte anhand einer statistischen Datenanalyse
- Schrittweiser Einsatz aller Objekte des Statistiklabors, vom Datenimport bis zum resultierenden Report.

Durch diese Gestaltung der Tutorials wird der Studierende sowohl an die Funktionalitäten der Objekte des Statistiklabors herangeführt als auch an eine statistisch sinnvolle Kombination dieser und statistisches Arbeiten.

Die Evaluationsstudie hat gezeigt, dass der Umgang mit dem Laborobjekt R-Kalkulator einen hohen Betreuungsaufwand erfordert und daher ein eigenes Tutorial für diesen benötigt wird. Das Tutorial sollte ausführlich gestaltet sein, da der R-Kalkulator eine entscheidende Komponente in Bezug auf die Aktivität der Studierenden darstellt. Ein Studierender, der den R-Kalkulator beherrscht, kann statistisch experimentieren und eigene Erfahrungen im Umgang mit Daten aufbauen. Das Tutorial sollte alle Besonderheiten des R-Kalkulators beachten und dessen Möglichkeiten zur Datenanalyse und strukturierten Programmierung umfassend vermitteln.

Die Erfahrungen der Evaluationsstudie führten zu einer kritischen Haltung des Autors gegenüber Tutorials. Es stellt sich die Frage, ob ein Tutorial die geforderten Leistungen erbringen kann. Zukünftige Forschungsprojekte sollten der Frage nachgehen, ob statistische Problemlösungskompetenz durch ein Tutorial transportiert werden kann oder ob es der direkten Betreuung der Teilnehmer bedarf.

---

<sup>319</sup>Traditionell im Bereich Statistik unterrichtete Studierende

Diese Überlegungen werden in neue Forschungsansätze eingebunden. Nach Abschluss des Projektes *Neue Statistik* wird über flächendeckende Einsätze des Statistiklabors nachgedacht. Der Entwicklung eines Tutorials wird dabei eine tragende Rolle zukommen. Die Auszeichnung des Statistiklabors mit dem Medienprix 2003 hat gezeigt, dass es sich gegenüber alternativen Entwicklungen auf diesem Gebiet erfolgreich durchsetzen konnte. CeDiS-Leiter Dr. Nicolas Apostolopoulos erklärt das prämierte Projekt:

„Das Statistiklabor ist eine explorative und interaktive Lernsoftware. Es fungiert als Werkzeugkasten zur Unterstützung der statistischen Ausbildung in allen Unterrichtsformen. Auf der Basis eines ausgefeilten didaktischen Konzepts<sup>320</sup> profitieren Lehrende und Studierende gleichermaßen von den Vorteilen der multimedialen und Computer gestützten Arbeitsumgebung. Einfache oder auch komplexe statistische Probleme können mit dem Statistiklabor in beherrschbare Details aufgeteilt und schrittweise individuell gelöst werden.“<sup>321</sup>

Es ist geplant, das Statistiklabor weiterzuentwickeln und flächendeckend einzusetzen, um einen Beitrag zur Qualität der Statistikausbildung an deutschen Hochschulen zu leisten.

---

<sup>320</sup>Hierzu ist kritisch zu bemerken, dass die Notwendigkeit zu einem solchen erst spät erkannt wurde. Die hier vorgestellten Komponenten des Einsatzkonzeptes stellen die einzigen Versuche in dieser Richtung dar.

<sup>321</sup>[Hertel2003], S. 15.

## Literatur

### I Bücher

- [AndersonLoynes1987] C.W. Anderson, R.M. Loynes: *The Teaching of Practical Statistics*, John Wiley & Sons, 1987
- [Beywl1988] W. Beywl: *Zur Weiterentwicklung in der Evaluationsmethodologie: Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation*, Verlag Peter Lang, Frankfurt am Main, 1988
- [BMBF1998] Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie: *Multimedia möglich machen. Deutschlands Weg in die Wissensgesellschaft*. Bonn, Februar 1998
- [Börner1992] W. Börner, G. Schnellhardt: *Multimedia. Grundlagen, Standards, Beispielanwendungen*, te-wi Verlag, München, 1992
- [Bowers1996] D. Bowers: *Statistics from Scratch: An Introduction for Health Care Professionals*, John Wiley & Sons, Chichester, 1996
- [Chatfield1995] Chr. Chatfield: *Problem Solving : a Statistician's Guide*, Chapman & Hall, London, 1995
- [DeckerWagner2002] R. Decker, R. Wagner: *Marketingforschung: Methoden und Modelle zur Bestimmung des Käuferverhaltens*, Verlag Moderne Industrie, München, 2002
- [DIALEKT2000] Digitale Interaktive Lektion, ein Forschungsvorhaben am Center für Digitale Systeme der Freien Universität Berlin: *Statistik Interaktiv*, Springer, 2000
- [Fahrmeir2001] L. Fahrmeir, R. Künstler, I. Pigoet, G. Tutz: *Statistik. Der Weg zur Datenanalyse*, 3. Auflage, Springer, Berlin, 2001
- [Fisher1925] R. A. Fisher: *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 1925
- [Fisher1956] R. A. Fisher: *Statistische Methoden für die Wissenschaft*, Oliver and Boyd, Edinburgh, 1956
- [Gigerenzer2003] G. Gigerenzer: *"Das Einmaleins der Skepsis", Über den richtigen Umgang mit Zahlen und Risiken*, 3. Auflage, Berlin Verlag, Berlin, 2003
- [Glaserfeld1995] H. v. Foerster, E. v. Glaserfeld, P. M. Hejl: *Einführung in den Konstruktivismus*, 2. Auflage, Piper, München, 1995
- [Glaserfeld1987] E. v. Glaserfeld: *Learning as a Constructive Activity*, in: Janvier, C. (Hrsg.): *Problems of Representation in the Teaching and Learning of Mathematics*, Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1987
- [GlowallaSchoop1992] U. Glowalla, E. Schoop: *Hypertext und Multimedia. Neue Wege in der computergestützten Aus- und Weiterbildung*, Springer Verlag, Heidelberg, 1992
- [Handl2002] A. Handl: *Multivariate Analysemethoden*, 1. Auflage, Springer Verlag, Heidelberg, 2002
- [Hauff1998] M. Hauff: *media-uni-multi.media? Entwicklung-Gestaltung-Evaluation neuer Medien*, Waxmann Verlag, Muenster, 1998
- [IssingKlimsma1997] L. J. Issing, P. Klimsma: *Informationen und Lernen mit Multimedia*. 2. überarbeitete Auflage, Beltz, Weinheim 1997

- [Jonassen1993] D. H. Jonassen: *Effects of semantically structured Hypertext Knowledge Bases*. In: C. A. Dillon and J. Richardson, Hrsg.: *Hypertext: A Psychological Perspective.*, Ellis Horwood, New York, 1993
- [Johnson1992] N. L. Johnson, S. Kotz, N. Balakrishnan: *Continuous univariate distributions*, 2. Auflage, Wiley, New York, 1992
- [Kerres2001] M. Kerres: *Multimediale und telemediale Lernumgebungen - Konzeption und Entwicklung*. 2., vollständig überarbeitete Auflage, Oldenbourg Verlag München Wien, 2001
- [Kindt1999] M. Kindt (Hrsg.): *Projektelevaluation in der Lehre, Multimedia an Hochschulen zeigt Profile*, Waxmann, Münster, 1999
- [Krämer1991] W. Krämer: *So lügt man mit Statistik*, Campus Verlag, Frankfurt/New York, 1991
- [Naeve2001] P. Naeve, *Statistik Interaktiv – Abschlußbericht WS 00/01*, Unveröffentlichtes Papier, Fakultät für Wirtschaftswissenschaften, Universität Bielefeld, 2001
- [Pausch1993] R. Pausch, F. E. Callies, *Interaktive Medien - Multimedia*, Ministerium für Wirtschaft, Landesinitiative Telekommunikation, Düsseldorf, 1993
- [Polya1990] G. Pólya: *How to solve it*, Penguin Books, St Ives, 1990
- [RiehmWingert1995] U. Riehm, B. Wingert: *Multimedia : Mythen, Chancen und Herausforderungen*, Bollmann, Mannheim 1995
- [RossMorrison2000] St. M. Ross, G. R. Morrison G.R.: *Evaluating interactive media: Issues and suggested practices*. In: Schenkel P., *Qualität und Evaluation von Lernsoftware*. BW Verlag für Bildung und Wissen, Nürnberg, 2000
- [Schenkel2000] Schenkel P., *Qualität und Evaluation von Lernsoftware. Evaluationsmethoden auf dem Prüfstand*, BW Verlag für Bildung und Wissen, Nürnberg, 2000
- [Schlittgen2000] R. Schlittgen: *Einführung in die Statistik, Analyse und Modellierung von Daten*, 9. Auflage, Oldenbourg, München, 2000
- [Schulmeister2002] R. Schulmeister: *Grundlagen hypermedialer Lernsysteme: Theorie-Didaktik-Design*, Addison-Wesley, Bonn, 2002
- [Scriven1996] M. Scriven: *Evaluation thesaurus. Introduction: The Nature of Evaluation*, 4. Auflage, Sage, Newbury Park, 1996
- [Simon1997] H. Simon (Hrsg.): *Virtueller Campus - Forschung für neues Lehren und Lernen*, Medien der Wissenschaft, Band 5, Waxmann-Verlag, Münster, 1997
- [Skinner1968] B.F. Skinner: *The Technology of Teaching*, Prentice Hall, New York, 1968
- [Sommerville2001] I. Sommerville: *Software Engineering*, 6. Auflage, Pearson Studium, München, 2001
- [Spitta1989] Th. Spitta: *Software Engineering und Prototyping*, Springer, Berlin, 1989
- [Stebler1994] R. Stebler, K. Reusser, C. Pauli: *Interaktive Lehr-Lern-Umgebungen: Didaktische Arrangements im Dienste des gründlichen Verstehens*. In K. Reusser und M. Reusser-Weyeneth: *Verstehen. Psychologischer Prozeß und didaktische Aufgabe*, Huber, Bern, 1994
- [Szekely1990] G. J. Szekely: *Paradoxa, Klassische und neue Überraschungen aus der Wahrscheinlichkeitsrechnung und mathematischer Statistik*, Harri Deutsch, Thun und Frankfurt am Main, 1990



- [Tiemann2002] V. Tiemann: *Kommunikationsprobleme in der Statistikausbildung*, Dissertation, Universität Bielefeld, Logos, Berlin, 2002
- [Weidenmann1986] B. Weidenmann: *Psychologie des Lernens mit Medien*, Psychologie Verlags Union, München, 1986
- [Worthen1988] B. R. Worthen: *Educational evaluation: Alternativ approaches and practical guidelines.*, Longman, London, 1988
- [Wottawa1998] H. Wottawa: *Lehrbuch Evaluation*, Verlag Hans Huber, Bern, 1998

## II Zeitschriften

- [Albert2003] J. H. Albert: *College Students' Conceptions of Probability*, The American Statistician, February 2003, Vol. 57, No. 1, 37 - 45
- [Bailey1998] R.A. Bailey: *Statistics and mathematics: The appropriate use of mathematics within statistics*, The Statistician 1998, 47, Parts2, pp. 261-271
- [Blackwell1986] D. Blackwell: *Conversation with David Blackwell*, Statistical Science, 1986, Vol. 1, No. 1, 40-53
- [Chan2002] G. Chan: *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*, The American Statistician, May 2002, Vol. 56, No.2, 156
- [Garfield2000] J. Garfield, B. Chance: *Assessment in Statistics Education: Issues and Challenges*, Mathematical Thinking and Learning, 2(1&2), 2000
- [Garfield1999] J. Garfield, I. Gal: *Assessment in Statistics Education: Current Challenges and Directions*, Mathematical Thinking and Learning, 2(1&2), 1999
- [Garfield1995] J. Garfield: *How Students Learn Statistics*, International Statistical Review, 1995, Vol. 63, Nr. 1, 25
- [Haunsperger1991] D. B. Haunsperger: *The Lack of Consistency for Statistical Decision Procedures*, The American Statistician, February, Vol. 45, No. 3, 1991, 252 - 255
- [Hertel2003] F. Hertel: *Lernsoftware macht das Rennen, Media Prix 2003 geht an die FU*, FU Nachrichten, Ausgabe 11/12 2003, S.15
- [HoaglinMoore1992] D. Hoaglin, D. Moore: *Perspectives on Contemporary Statistics*, MAA Notes, 21, 1992, 1 - 17
- [Kendall1968] M. G. Kendall: *On the future of statistics-a second look (with discussion)*, J.Roy.Statist.Soc. A. 131, 182-204
- [Moore1995] D. S. Moore, G. W. Cobb, J. Garfield, W. Q. Meeker: *Statistics Education Fin de Siècle*, The American Statistician, 1995, Vol. 49, No. 3, 250 - 260
- [Moore1997] D. S. Moore: *New Pedagogy and New Content: The Case of Statistics*, International Statistical Review, 1997, Vol. 65, No. 2, 123 - 165
- [Perkins1999] D. Perkins: *The Many Faces of Constructivism*, Educational Leadership, November, 1999
- [Sawyer2002] T. Sawyer, Ohio Congressman: *Presidential Invited Address given at the Joint Statistical Meetings in New York on August 12, 2002*, Amstat News, November 2002, 8 -13
- [Scheaffer2002] R. L. Scheaffer: *Statistical Bridges*, Journal of the American Statistical Association, March 2002, Vol. 97, No 457, 1 - 7

- [Straf2002] M. Straf: *Pathway to Statistics*, Amstat News, July 2002
- [Straf2002] M. Straf: *What is Statistics?*, Amstat News, Oktober 2002, 2 - 3
- [Symanzik2003] J. Symanzik: *Teaching Experiences with a course on "Web-Based Statistics"*, The American Statistician, February 2003, Vol. 57, No. 1, 46 - 50
- [Thomas2003] R. Thomas: *Pstatistics and Mstatistics*, RSS News Leitartel, Mai 2003, Vol, 30, Nr. 9
- [Utts2003] J. Utts: *What educated citizens should know about Statistics and Probability*, The American Statistician, May 2003, Vol. 57, No. 2, 74 - 79
- [VardemanMorris2003] St. B. Vardeman, M. D. Morris: *Statistics and Ethics: Some Advice for Young Statisticians*, The American Statistician, Februar 2003, Vol. 57, No. 1, 21 - 26
- [Vanderpump1996] M. P. Vanderpump: *Ignoring a Covariate: An Example of Simpson's Paradox*, The American Statistician, February 1996, Vol. 50, No. 4, 340 -341
- [Wardrop1995] R. L. Wardrop: *Simpson's Paradox and the Hot Hand in Basketball*, The American Statistician, February 1995, Vol. 49, No. 1, 24 - 28
- [Wagner1982] C. H. Wagner: *Simpson's Paradox in Real Life*, The American Statistician, Februar 1982, Vol. 36, Nr. 1, 46 - 49

### III Internetquellen

- [Roosevelt] <http://home.t-online.de/home/hartmut.weist/statistik.htm>
- [SimpsonRichmond] [http://mitglied.lycos.de/paradoxon\\_2/english/paradoxes.html](http://mitglied.lycos.de/paradoxon_2/english/paradoxes.html)
- [Wardrop2003] <http://www.stat.wisc.edu/~wardrop/courses/302ch52002.pdf>, Author: Robert L. Wardrop Professor Department of Statistics University of Wisconsin - Madison
- [Murphy1997] Elizabeth Murphy: "Constructivism - From Philosophy to Practice", University Laval, Quebec, Summer 1997 with Jacques Rheaume. Webpage, created in context of course TEN-62349, <http://www.stemnet.nf.ca/el-murphy/emurphy/cle.html>

# Materialien

## A CD-Rom: Statistiklabor, Laborseiten, Bibliotheken

Auf der beigefügten CD-ROM befinden sich folgende Ordner:

- Allgemeine Bibliothek: Hier befindet sich die Bibliothek *Datenanalyse.R*.
- Evaluationsinstrument: Hier befinden sich alle Laborseiten des in Teil IV demonstrierten Evaluationsinstrumentes sowie der Evaluationsreport.
- LernmodulVar enthält die Theoriekomponente (erstellt mit dem Lernmoduleditor), die in I.4.4.1. vorgestellt wurde.
- Readme enthält Benutzeranweisungen zum Umgang mit Statistiklaborseiten.
- SimpsonParadox, Stichprobengröße, Strukturen und Varianzzerlegung enthalten die Laborseiten zu I.4.
- Statistiklabor enthält die Version 2.0.1.4 des Statistiklabors. Diese ist im Forschungsprojekt *Neue Statistik* entstanden.

## B Benutzeranweisungen zum Umgang mit Statistiklaborseiten

Im Teil I der Dissertation werden vier Beispiele gegeben, statistische Bereiche mit Hilfe des Statistiklabors zu vermitteln. Die für diese Zwecke erstellten Laborseiten (Endung: spf) stehen dem Leser auf der beigefügten CD-ROM<sup>322</sup> bereit. Diese enthält die folgenden Gebiete und Laborseiten, die auch durch die Ordnerstruktur erkennbar sind:

- Varianzzerlegung: Maschinenoutput, Schuhe, Schuhe2, Schuhe3
- Simpson-Paradox: Simpson, Simpson3x2
- Stichprobengröße: Erhebung1, Erhebung2, Erhebung3, Erhebung4, Erhebung5, Erhebung6, Erhebung7, Erhebung8
- Strukturen: Transformation1, Transformation2, Transformation3, Transformation4, Transformation5, Transformation6, Transformation7, Transformation8, Transformation9, Transformation10

Für jeden der vier Bereiche existieren Bibliotheken, die den R-Code enthalten, der auf den entsprechenden Laborseiten benötigt wird. Die R-Kalkulatoren der Laborseiten beinhalten Funktionsaufrufe. Diese rufen die Funktionen der entsprechenden Bibliothek auf. Es erfolgt eine Zuordnung der erstellten Bibliotheken (Endung: R) je Bereich:

- Varianzzerlegung: maschinen, schuhe, schuhe2, schuhe3
- Simpson-Paradox: ktsimpson
- Stichprobengröße: stichprobengröße
- Strukturen: strukturen

Für die Funktionsfähigkeit der Laborseiten müssen die Bibliotheken eingebunden werden. Dies erfolgt nach dem Öffnen einer Laborseite auf der Taskleiste unter den aufeinander folgenden Menüpunkten: *Projekt-Bibliotheken-Bibliotheken einbinden*. Auf jeder Laborseite befindet sich das Objekt R-Kalkulator. In diesem finden alle R-Operationen statt, z.B. die Funktionsaufrufe. Der Benutzer bekommt in dem Objekt *Texteditor* Ziele und Umgang mit dem Inhalt des R-Kalkulators mitgeteilt. Der R-Kalkulator kann sich in zwei verschiedenen Zuständen/Modi befinden: dem Run-

---

<sup>322</sup>Vgl. Anhang A.

oder dem Edit-Modus. Zwischen beiden kann auf der Laborseite variiert werden: der Run-Modus ist aktiv, wenn das graue Rad angeklickt wurde, der Edit-Modus nach Anklicken des Stiftsymbols. Im Edit-Modus kann der Funktionsaufruf bzw. R-Code eingegeben werden. Im Run-Modus wird das Programm ausgeführt.

Es gibt zwei Typen von Statistiklaborseiten: Demonstrations- und Interaktionsseiten. Sofern die Möglichkeit der Interaktivität besteht, wird dies dem Benutzer im Texteditor auf der entsprechenden Laborseite mitgeteilt. Da der R-Code jeder einzelnen Bibliothek innerhalb dieser dokumentiert ist, besteht die Möglichkeit, jede (auch die Demonstrationsseiten) zu modifizieren. Dies ist ein Angebot an den Dozenten, der die Laborseiten individuell nutzen möchte.

Um das Statistiklabor zu installieren, kann der Benutzer die beigefügte CD-Rom (Version 2.0.1.4) benutzen.

Die Laborseiten wurden nach dem sogenannten Schichtenprinzip gestaltet. Die Studierenden sehen in dem R-Kalkulator jeder Laborseite so wenig R-Code wie möglich. Ein Großteil der Laborseiten enthält lediglich den Funktionsaufruf im R-Kalkulator, der gesamte R-Code befindet sich eine Ebene tiefer in der zugehörigen Bibliothek. Im Entwicklungsprozess der Laborseiten gab es drei Stufen:

1. Der gesamte R-Code befand sich im R-Kalkulator und war für den Benutzer sofort sichtbar.
2. Der R-Code befand sich weiterhin im R-Kalkulator, war aber nicht sofort sichtbar. Der R-Kalkulator wurde auf Symbolansicht gestellt.
3. Der gesamte R-Code wurde in eine Bibliothek ausgelagert, die für den Studierenden nicht sichtbar ist. Der R-Kalkulator enthält lediglich den Funktionsaufruf. Ein Editor erklärt auf der Laborseite Ziele und Umgang mit dem R-Kalkulator.

Das Ziel bestand darin, den R-Code nicht sichtbar für den Studierenden zu machen (Prinzip des Informationhiding), dem Dozenten aber jederzeit zur Verfügung zu stellen (Transparenz, Modifizierbarkeit).

## C Websites für Updates und weitergehende Informationen

Auf der Webpräsenz des Forschungsprojektes *Neue Statistik* hat der Leser die Möglichkeit, sich Updates des Statistiklabors und weitere Materialien wie z.B. Applets und Animationen downzuloaden. Die Website [www.neuestatistik.de](http://www.neuestatistik.de) ist Passwort geschützt, um die Materialien vor Missbrauch zu schützen. Der Leser kann sich jederzeit an den Autor wenden, um ein Benutzerlogin für den wissenschaftlichen Gebrauch zu erhalten.

## D 1. Evaluationsbogen - Semesterbeginn

Liebe Studentinnen und Studenten,

Im Rahmen des Forschungsprojektes NEUE STATISTIK testen wir die multimediale Lernsoftware *Statistik interaktiv komplett* sowie ihren Einsatz im Lernmanagementsystem Blackboard.

Wir sind auf Ihre aktive Hilfe angewiesen. Bitte füllen Sie den folgenden Fragebogen gewissenhaft aus. Besten Dank im Voraus. Selbstverständlich werden alle Ihre Angaben anonym behandelt. Für Rückfragen steht Ihnen Frau Pabst oder Herr Nölke gern zur Verfügung.

Selbstverständlich werden alle von Ihnen gemachten Angaben **vertraulich** behandelt und nicht an Dritte weitergegeben.

1. Welche der folgenden Kommunikationsmöglichkeiten stehen Ihnen privat zur Verfügung?  $\hookrightarrow$  **Mehrfachnennungen möglich**  $\hookrightarrow$

- Handy  Zugang zu Telefonanschluss  
 Analoges Modem  Communicator  
 DSL

2. Welche der folgenden Mediennutzungsmöglichkeiten stehen Ihnen privat zur Verfügung?  $\hookrightarrow$  **Mehrfachnennungen möglich**  $\hookrightarrow$

- Palmtop  Eigenes Notebook  
 Eigener PC  Pocket-PC  
 Apple-Macintosh

3. Welche der folgenden technischen Ausstattungen stehen Ihnen privat zur Verfügung?  $\hookrightarrow$  **Mehrfachnennungen möglich**  $\hookrightarrow$

- CD-Brenner  DVD-Brenner  
 MP3-Player  TFT-Flachbildschirm  
 DVD-Player

4. Wie lange surfen Sie pro Woche durchschnittlich im Internet?

- gar nicht  weniger als 30 Min  
 weniger als 1 Std  weniger als 2 Std  
 weniger als 5 Std  weniger als 7 Std  
 7 - 10 Std  mehr als 10 Std

5. Ich nutze das Internet beruflich in folgender wöchentlicher durchschnittlicher Nutzungsdauer:

- gar nicht  weniger als 30 Min  
 weniger als 1 Std  weniger als 2 Std  
 weniger als 5 Std  weniger als 7 Std  
 7 - 10 Std  mehr als 10 Std

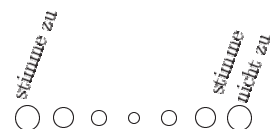
6. Wie lange nutzen Sie durchschnittlich das Internet privat pro Woche?

- gar nicht  weniger als 30 Min  
 weniger als 1 Std  weniger als 2 Std  
 weniger als 5 Std  weniger als 7 Std  
 7 - 10 Std  mehr als 10 Std

7. Geben Sie in 3 Sätzen Ihre bisherige Vorstellung von Statistik an.

$\rightarrow$  \_\_\_\_\_  
 $\rightarrow$  \_\_\_\_\_  
 $\rightarrow$  \_\_\_\_\_

8. Nehmen Sie bitte Stellung zu der folgenden Aussage. Der Einsatz der Multimediasoftware Statistik-Interaktiv-komplett ermöglicht eine realitätsnahe Vermittlung der Statistik:



9. Wie wichtig erscheint Ihnen eine realitätsnahe und aktuelle Vermittlung der Statistik durch die Multimediasoftware Statistik-Interaktiv-komplett?



10. Halten Sie die Statistik für eine fächerübergreifende Disziplin?

- ja  nein

11. Wie wichtig erachten Sie statistische Methoden in den folgenden Bereichen:

	wichtig				unwichtig
Marketing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Controlling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Finanzwirtschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Produktionswirtschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Personal und Organisation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Betriebliche Steuerlehre	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12. Haben Sie Übungsaufgaben im Statistiklabor gelöst, die nicht Teil der Übung waren?

ja       nein

13. Haben Sie multimediale Komponenten betrachtet, die nicht vom Dozenten in der Vorlesung empfohlen oder bearbeitet wurden?

ja       nein

14. Haben Sie sich an Vorlesung und Übung aktiv beteiligt?

ja       nein

15. Haben Sie Laborobjekte ausprobiert, die bisher nicht in der Veranstaltung relevant waren?

ja       nein

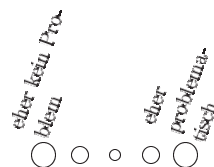
16. Haben Sie das Statistiklabor angewendet, um Probleme ausserhalb der Veranstaltung zu lösen?:

- Ja, und zwar in den folgenden Bereichen:
  - Job
  - VWL
  - Methoden
  - privat
  - BWL
  - Andere: \_\_\_\_\_
- Nein

17. Haben Sie vor dieser Veranstaltung mit einer Programmiersprache gearbeitet?

ja       nein

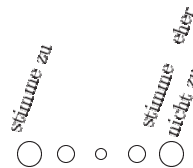
18. Wie schätzen Sie Ihre Fähigkeit ein, ein kleines gegebenes aber unbekanntes Problem nur mit Hilfe des Statistiklabors zu lösen?



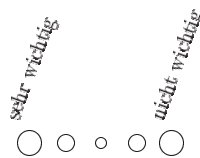
19. Welches Labor-Objekt bevorzugen Sie bei einer Lösungserstellung?

- R-Grafik-Wizzard
- R-Kalkulator
- eher andere, und zwar \_\_\_\_\_

20. Nehmen Sie Stellung zu folgender Aussage: Der R-Kalkulator stellt ein elementares Mittel zur statistischen Problemlösung dar.



21. Wie wichtig erachten Sie die Erlernung der Anwendung des R-Kalkulators?



22. Haben Sie bereitgestellte Hilfen im Statistiklabor (Hilfe und Online-Tutorial) genutzt?

- Ja, und zwar für die folgenden Objekte:
  - R-Kalkulator
  - Grafikwizzard
  - Datensatzimport
  - R-Grafik
  - Datensatzexport
- Nein

23. Wie schätzen Sie den Einfluss des Statistiklabors auf Ihre später notwendigen statistisch-methodischen Fähigkeiten ein?

sehr hilfreich  
 ○ ○ ○ ○ ○  
 gar nicht hilf.  
 reich

24. Halten Sie die Lernsoftware Statistik Interaktiv komplett als notwendig für die statistischen Vorlesungsinhalte?

ja            nein  
           

25. Wie wichtig erscheint Ihnen der Einsatz des Lernmanagementsystems Blackboard?

sehr wichtig  
 ○ ○ ○ ○ ○  
 unwichtig

26. Nehmen Sie Stellung zu folgender These:  
Die Lernsoftware Statistik Interaktiv komplett erleichtert mir das Erlernen der Statistik.

trifft zu  
 ○ ○ ○ ○ ○  
 trifft nicht zu

27. Wie wichtig bewerten Sie das Ziel, im Rahmen Ihrer Ausbildung statistische Grundkenntnisse zu erwerben?

sehr wichtig  
 ○ ○ ○ ○ ○  
 unwichtig

28. Trauen Sie sich zu mit allen Komponenten der Lernsoftware Statistik Interaktiv komplett die Vorlesung nachzuarbeiten?

ja            nein  
           

29. Wie schätzen Sie Ihre Fähigkeit ein, nach der Erstellung einer Laborseite in Vorlesung oder Übung das Labor eigenständig zu nutzen?

kein Problem  
 ○ ○ ○ ○ ○  
 unmöglich

30. Haben Sie schon selbstständig Laborseiten erstellt?

ja            nein  
           

31. Haben Sie aufgrund Ihrer bisherigen Kenntnisse in Statistik und ersten Erfahrungen im Statistiklabor den Eindruck, dass Ihnen dieses das statistische Arbeiten erleichtert?

ja            nein  
           

32. Haben Sie aufgrund Ihrer bisherigen Kenntnisse in Statistik und ersten Erfahrungen im Statistiklabor den Eindruck, dass das Statistiklabor umfassende statistische Lösungsmöglichkeiten bietet?

ja            nein  
           

33. Wie beurteilen Sie Ihren bisherigen Umgang/Kontakt mit Neuen Medien?

sehr umfassend  
 ○ ○ ○ ○ ○  
 noch gar nicht

34. Ermöglicht Ihnen das Lernmanagementsystem Blackboard eine bessere Organisation Ihres Lernens?

ja      nein  
     

35. Nehmen Sie Stellung zu folgender These:  
 Das Lernmanagementsystem Blackboard ist ideal um sämtliche Veranstaltungskomponenten miteinander zu vereinen.

trifft zu      trifft nicht zu

36. Nehmen Sie Stellung zu folgender These:  
 Das Lernmanagementsystem Blackboard stellt eine praktische Plattform zur Kommunikation/zum Austausch zwischen Studierenden untereinander und zu Dozenten dar (z.B. Report).

trifft zu      trifft nicht zu

37. Nehmen Sie Stellung zu folgender These:  
 Das Lernmanagementsystem Blackboard ist überflüssig, da durch eine Kurs-Web-Seite und der Austausch über e-mail gleiche Ziele erreicht werden können.

trifft zu      trifft nicht zu

38. Wie stark interessieren Sie sich für das Fach Statistik?

Gar nicht      Sehr stark      Weiss nicht

39. Wie schwierig schätzen Sie das Fach Statistik **allgemein** ein?

Nicht schwierig      Sehr schwierig      Weiss nicht

40. Wie schwierig schätzen Sie das Fach Statistik **für Sie persönlich** ein?

Nicht schwierig      Sehr schwierig      Weiss nicht

41. Wie wichtig ist das Fach Statistik Ihrer Meinung nach für Ihren Studiengang?

Nicht wichtig      Sehr wichtig      Weiss nicht

42. In welchem Ausmaß brauchen Sie Ihrer Meinung nach das Fach Statistik in Ihrem **späteren Berufsleben**?

Gar nicht      Sehr stark      Weiss nicht



43. Wie wichtig sind Ihnen die folgenden Eigenschaften des Lernprogramms?

			sehr wichtig						sehr unwichtig
Inhaltlicher Aufbau nachvollziehbar			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beispiele des vermittelten Stoffes			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angemessenes Lerntempo			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aktive Mitarbeit mit dem Programm			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anspruchsvolle thematische Inhalte			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Komplizierte Sachverhalte einfach darstellen			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhalte nützlich für das Studium			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Weckung des Interesses an Statistik			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

44. Wie zutreffend sind die folgenden Eigenschaften für das Statistik Lernprogramm Ihrer Meinung nach?

			sehr wichtig						sehr unwichtig
Inhaltlicher Aufbau nachvollziehbar			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beispiele des vermittelten Stoffes			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angemessenes Lerntempo			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aktive Mitarbeit mit dem Programm			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anspruchsvolle thematische Inhalte			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Komplizierte Sachverhalte einfach darstellen			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhalte nützlich für das Studium			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Weckung des Interesses an Statistik			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Nehmen Sie bitte Stellung zu folgender Aussagen:

45. Die Inhalte der Lernumgebung sind gut strukturiert.

Trifft gar nicht zu	Trifft voll- kommen zu
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

46. Nehmen Sie bitte Stellung zu folgender Aussage:  
Die Nutzung der Lernumgebung ist für Neulinge einfach.

Trifft gar nicht zu	Trifft voll- kommen zu
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

47. Nehmen Sie bitte Stellung zu folgender Aussage:  
Ich würde die Lernumgebung weiterempfehlen.

Trifft gar nicht zu	Trifft voll- kommen zu
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

48. Nehmen Sie bitte Stellung zu folgender Aussage:  
Ich habe den Stoff regelmäßig vor- und nachgearbeitet.

Trifft gar nicht zu	Trifft voll- kommen zu
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

49. Wie haben Sie die Lernumgebung am häufigsten genutzt?

In der Vorlesung	In der Übung	Zu Hause	In der Recherstunde
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vielen Dank für Ihre Mitarbeit!

## E 2. Evaluationsbogen - Semesterende

Liebe Studentinnen und Studenten,

Im Rahmen des Forschungsprojektes NEUE STATISTIK testen wir die multimediale Lernsoftware *Statistik interaktiv komplett* sowie ihren Einsatz im Lernmanagementsystem Blackboard.

Wir sind auf Ihre aktive Hilfe angewiesen. Bitte füllen Sie den folgenden Fragebogen gewissenhaft aus. Besten Dank im Voraus. Selbstverständlich werden alle Ihre Angaben anonym behandelt. Für Rückfragen steht Ihnen Frau Pabst oder Herr Nölke gern zur Verfügung.

Selbstverständlich werden alle von Ihnen gemachten Angaben **vertraulich** behandelt und nicht an Dritte weitergegeben.

1. Nehmen Sie bitte Stellung zu der folgenden Aussage.  
Der Einsatz der Multimediasoftware Statistik-Interaktiv-komplett ermöglicht eine realitätsnahe Vermittlung der Statistik:

*stimme zu*                         *stimme nicht zu*

2. Wie wichtig erscheint Ihnen eine realitätsnahe und aktuelle Vermittlung der Statistik durch die Multimediasoftware Statistik-Interaktiv-komplett?

*wichtig*                     *unwichtig*

3. Halten Sie die Statistik für eine fächerübergreifende Disziplin?

ja     nein

4. Nehmen Sie Stellung zu folgender Aussage:  
Statistisches Reportwriting kann im Team besser erlernt werden als in Einzelarbeit.

*stimme zu*             *stimme nicht zu* <sup>etwa</sup>

5. Nehmen Sie Stellung zu folgender Aussage:  
Die kontinuierliche Bearbeitung von Aufgaben im Statistikkolabor verbesserte meinen Umgang mit dem Statistikkolabor.

*stimme zu*             *stimme nicht zu* <sup>etwa</sup>

6. Wie schätzen Sie die Bearbeitung der ersten Aufgaben im Labor im Bezug auf die Vorbereitung zur Erstellung der ersten Mini-Reports ein?

*sehr hilfreich*             *nicht hilfreich*

7. Wie schätzen Sie die Erstellung der ersten Mini-Reports im Bezug auf die Vorbereitung zur Erstellung der ersten Reports ein?

*sehr hilfreich*             *nicht hilfreich*

8. Wie schätzen Sie die Bearbeitung des ersten Reports im Bezug auf die Vorbereitung zur Erstellung der dritten Reports (Einzelleistung) ein?

*sehr hilfreich*             *nicht hilfreich*

9. Haben Sie das Statistikkolabor angewendet, um Probleme außerhalb der Veranstaltung zu lösen?:

Ja, und zwar in den folgenden Bereichen:

- Job     privat  
 VWL     BWL  
 Methoden     Andere: \_\_\_\_\_

Nein

10. Nehmen Sie Stellung zu folgender Aussage:  
 Der R-Kalkulator stellt ein elementares Mittel zur statistischen Problemlösung dar.

stimme zu eher  
 nicht zu

---

11. Wie wichtig erachten Sie die Erlernung der Anwendung des R-Kalkulators?

sehr wichtig nicht wichtig

---

12. Nehmen Sie Stellung zu folgender These:  
 Die Lernsoftware Statistik Interaktiv komplett erleichtert mir das Erlernen der Statistik.

trifft zu trifft nicht zu

---

13. Wie wichtig bewerten Sie das Ziel, im Rahmen Ihrer Ausbildung statistische Grundkenntnisse zu erwerben?

sehr wichtig unwichtig

---

14. Wie schätzen Sie Ihre Fähigkeit ein, nach der Erstellung einer Laborseite in Vorlesung oder Übung das Labor eigenständig zu nutzen?

kein Problem unmöglich

15. Trauen Sie sich zu, mit allen Komponenten der Lernsoftware Statistik Interaktiv komplett die Vorlesung nachzuarbeiten?

ja nein

---

16. Nehmen Sie Stellung zu folgender These:  
 Das Lernmanagementsystem Blackboard ist ideal, um sämtliche Veranstaltungskomponenten miteinander zu vereinen.

trifft zu trifft nicht zu

---

17. Nehmen Sie Stellung zu folgender These:  
 Das Lernmanagementsystem Blackboard stellt eine praktische Plattform zur Kommunikation/zum Austausch zwischen Studierenden untereinander und zu Dozenten dar (z.B. Report).

trifft zu trifft nicht zu

---

18. Konnte Ihr Interesse für Statistik durch das Projekt geweckt werden? Wenn ja, ist dies eher auf den Einsatz von Multimedia zurückzuführen oder auf andere Aspekte wie zum Beispiel die Betreuung, Form der Vorlesung, ...

---



---



---



---



---

19. *Empfanden Sie es als positiv, Teile der Prüfungsleistung schon während des Semesters erbringen zu können (Aufgaben, Reports) - wenn ja, warum? ... oder empfanden Sie die höhere Belastung während des Semesters als eher störend (Gründe?)?*

---

---

---

---

---

---

---

20. *Das Projekt besteht aus verschiedenen Komponenten: Vorlesung, Übung, Rechnerstunden, selbstständiges Reportwriting - welche Meinung haben Sie bezüglich der einzelnen Komponenten (Stichworte), wo haben Sie am meisten gelernt? Verbesserungsvorschläge/Kritik?*

---

---

---

---

---

---

---

21. *Würden Sie sich wiederholt für ein Multimediaprojekt anmelden, wenn Sie die Möglichkeit dazu hätten? (Begründung)*

---

---

---

---

---

---

---

22. *Hat Ihnen das Reportwriting das wissenschaftliche Arbeiten nähergebracht?*

---

---

---

---

---

23. *Beurteilen Sie die einzelnen Komponenten der eingesetzten Multimediasoftware.*

---

---

---

---

---

---

---

24. *Welche der in der Vorlesung vermittelten statistischen Gebiete finden Sie am interessantesten, z.B. weil Sie sich vorstellen können, sie später noch gebrauchen zu können?*

---

---

---

---

---

---

---

25. *Wie schätzen sie das Teamwork ein? (Begründung)*

---

---

---

---

---

---

---

---

26. *Sehen Sie die Erstellung der statistischen Reports als gute Einführung in das statistische Problemlösen? (Begründung)*

---

---

---

---

---

---

## F Qualitative Befragung der 1. Durchführung

*Statistik interaktiv!* WS 2001/2002 Evaluationsbogen zum Semesterende

1. **Konnte Ihr Interesse für Statistik durch das Projekt geweckt werden? Wenn ja, ist dies eher auf den Einsatz von Multimedia zurückzuführen oder auf andere Aspekte wie zum Beispiel die Betreuung, Form der Vorlesung, ...**

Das Interesse der Studierenden wurde vor allem durch den Einsatz von Multimedia und die Arbeit in der kleinen Gruppe geweckt. Sie betonten immer wieder die gute Kommunikation mit den Dozenten: „Man hat gemerkt, dass sich um einen gekümmert wurde.“ Statistik in Zusammenhang mit der Realität kennenzulernen empfand ein Großteil der Studierenden als sehr interessant, besser als die mathematische Seite der Statistik. „Die ermittelten Ergebnisse sinnvoll zu interpretieren, ist für mich wesentlich wichtiger als stundenlang ein Ergebnis auszurechnen und dann nicht zu wissen, was diese Zahl bedeutet.“ „Im Gegensatz zu allen anderen Fächern hat man hier auch mal Probleme aus der Praxis erlebt und nicht nur stupide trockene Theorie gepaukt.“

Kommunikation

Auch wurde das Interesse durch die Gestaltung der Veranstaltung geweckt (Motivation, Theorie, Aufgaben). Das Interesse der Studierenden blieb über das ganze Semester hinweg bestehen, da sie sich ständig selbst anhand der Aufgaben mit der Materie auseinandersetzen. Es wurde auch immer wieder erwähnt, dass die realitätsnahen Reports zum Lernen motivierten. „Aktivität, selbstständiges Denken und Verstehen wurde viel stärker gefordert und gefördert als in den anderen Grundstudiumsveranstaltungen. Dazu hat der Einsatz von Multimedia beigetragen, weil z.B. Reports in diesem Umfang möglich wurden, aber auch die kleine Gruppe.“

Praxisbezug

Als spannend und interessant wurden der deutliche Bezug zum täglichen Leben sowie Raum und Zeit zur Diskussion untereinander sowie mit den Dozenten bezeichnet. Es wurde deutlich, dass dabei die intensive Betreuung sehr hoch einstuft wird. Viele Studierende bemerkten, dass sie einen allgemeinen Eindruck von der Statistik gewonnen haben. „Mir ist klar geworden, wie viel man täglich mit der Statistik zu tun hat und ich achte jetzt auch vermehrt auf statistische Analysen.“ ... „meiner Meinung nach hat das Programm *Statistik interaktiv* den Einstieg in die Theorie sehr vereinfacht. Man drückt auf einen Button und schon ist das Histogramm erstellt.“ ... „vielleicht wird einem die Angst vor der Erstellung theoretischer Analysen ein wenig genommen, weil das multimedial doch so einfach sein kann.“

Kleine Gruppe

Ein Nachrücker, der die ersten Wochen über die traditionelle Veranstaltung besuchte, äußerte sich wie folgt: „.. Statistik war halt nur ein weiteres theoretisches Fach in meinem Studienplan. ... In der Folgezeit änderte sich meine Auffassung ... Das ist sicherlich auf die Anwendung von Multimedia zurückzuführen“. Viele Studierende betonten, dass ihr Interesse an Statistik besonders durch das Hantieren mit Datensätzen im Labor wuchs. Doch es gab auch andere Stimmen. „Da das Programm nur anfangs auf meinem Computer lief (nicht XP tauglich), konnte ich von dem Einsatz von Multimedia nicht sehr profitieren.“

Lob auf Multimedia

Aus dem Großteil der Antworten geht deutlich hervor, dass das Interesse primär durch die Form der Veranstaltung und sekundär durch Multimedia geweckt wurde. „Das Projekt hat mein Interesse für Statistik verstärkt, allerdings war der Multimediaeinsatz nebensächlich. Entscheidend war, dass die Vorlesung in einer kleinen Gruppe stattfand.“ „ Die übersichtlich präsentierte Theorie und abwechslungsreiche Gestaltung der Vorlesungen durch Multimedia weckte Interesse.“

2. **Empfanden Sie es als positiv, Teile der Prüfungsleistung schon während des Semesters erbringen zu können (Aufgaben, Reports) - wenn ja, warum? ... Oder empfanden Sie die höhere Belastung während des Semesters als eher störend (Gründe)?**

Generell äußerten sich die Studierenden sehr positiv darüber, kontinuierlich Prüfungsleistungen zu erbringen. „ ... und nicht durch eine einmalige Hammerklausur, die niemals das tatsächliche Können und Wissen der Studierenden widerspiegeln kann.“ Allerdings wurden

die wöchentlichen Aufgabenzettel als zu schwer und umfangreich bezeichnet. ... „teilweise zu viel abverlangt, folglich haben wir zeitweise über 20 h in der Woche nur an Statistik gegessen.“ Besonders in der Zeit direkt vor den Klausuren, d.h. innerhalb der letzten Wochen des Semesters empfand ein Großteil der Studierenden die wöchentliche Belastung durch die Aufgaben als belastend, da der Anspruch der Aufgabenblätter zunahm. Positiv wurde immer wieder die ständige Überarbeitung des laufenden Stoffes durch Aufgabenabgabepflicht hervorgehoben.

Hohe Belastung

Auch die Unterteilung der Vorlesung in Motivation und Theorie wurde als sehr positiv empfunden. Auch in diesem Zusammenhang wurde der Zwang zum Nacharbeiten des Stoffes im Nachhinein sehr positiv bewertet. Durch ständiges Verstehen fühlte sich ein Großteil motiviert, mehr von der Statistik zu lernen.

Dass sie durch die kontinuierlichen Arbeiten während des Semesters am Semesterende eine Klausur weniger hatten, wurde als großer Vorteil herausgestellt. Auch die statistischen Reports während des Semesters wurden als positiver Druck empfunden, sich ernsthaft mit den statistischen Methoden auseinanderzusetzen. „Die Idee, die Prüfungsleistungen schon während des Semesters zu erbringen, fand ich klasse. Man ist sozusagen gezwungen, kontinuierlich mitzuarbeiten, und kann somit den Faden nicht verlieren.“ Positiv wurde immer wieder die Möglichkeit bewertet, durch die Prüfungsleistungen während des Semesters eine evtl. schlechte Klausurleistung (durch z.B. Blackout) ausgleichen zu können. Die Aufgaben und Reports bezeichneten viele als positive Herausforderung im Semester und empfanden die Entlastung durch eine kurze Multiplechoiceklausur als guten Kompromiss. „Während des Semesters viele Aufgaben zu machen und ein paar Reports zu schreiben, hat mich manchmal schon sehr belastet, aber ich halte dies für hervorragend. Druck kann auch sehr positiv sein!“ Der Vorteil, wöchentlich das erlernte Wissen anzuwenden und zu vertiefen, wird immer wieder zusammen mit dem hohen Zeitaufwand genannt. „Wenn man nun letztendlich den wöchentlichen Zeitaufwand für Statistik betrachtet, kommt die Frage auf, ob man BWL oder Statistik studiert!“

positiver Druck

hoher Zeitaufwand

### 3. Das Projekt besteht aus verschiedenen Komponenten: Motivationsvorlesung, Theorievorlesung, Tutorium, selbstständiges Reportverfassen - welche Meinung haben Sie bezüglich der einzelnen Komponenten (kurz, Stichworte), wo haben Sie am meisten gelernt?

**Kommentare zur Motivationsvorlesung:** Die Studierenden fanden es sehr sinnvoll, den Zusammenhang zwischen der Statistik und dem Leben aufgezeigt zu bekommen. „... sehr hilfreich und teilweise auch dringend notwendig, um die Theorievorlesung in den praktischen Kontext zu bringen ... aber teilweise zu ausführlich und langatmig.“ Doch auch andere Stimmen wurden laut: „... war ab und an ganz nett, sich bei einem Kaffee ein paar statistische Geschichten anzuhören, aber reichlich früh“; „wirkte auf mich sehr verwirrend“. Positiv wurde bemerkt, dass man ohne schwere Mathematik an die Statistik herangeführt wurde. Dies wurde von vielen als angenehmer Einstieg in den Stoff angesehen. „... etwas verwirrend, jedoch gab sie gute Einblicke in die praktische Anwendung von Statistik.“

Zusammenhänge

**Kommentare zur Theorievorlesung:** Sie wurde von einem Großteil als die Komponente angesehen, in der am meisten gelernt wurde. Die Stoffvermittlung wurde jedoch als zu schnell und zu umfangreich beurteilt. „Die Mitschriften sind zwar super, aber ansatzweise den Sinn verstehen konnte man erst zu Hause, und dann war es zu spät zum Fragen.“

zu viel zu schnell

Viele betonten, dass sie die klar und gut strukturierte Vorlesung benötigten, um den Stoff der Motivationsveranstaltung erklärt zu bekommen; „Die Theorievorlesungen haben sehr anschaulich die Mathematik vermittelt, außerdem konnte man ihnen ein Skript abgewinnen, das nützlich zur Lösung der Aufgaben war.“ Von vielen Studierenden wurde das Skript deshalb positiv bewertet, weil es sich aufs Wesentliche beschränkte. „Dank der kurzen und sinnvollen Zusammenfassungen wurde man von Informationen nicht so überschwemmt wie in den Lehrbüchern.“

gute Struktur

**Kommentare zum Tutorium:** Es gab ausnahmslos sehr schlechte Kritik mit den folgenden Begründungen: Keine Verbesserung und Erklärung von Fehlern, keine Erklärung zu den Lösungen, schlechte Korrektur der Aufgaben, keine Anwesenheit in den Rechnerstunden, bei Fragen Verweis auf Dozenten oder: das ist so! - „Hatte man die Aufgaben richtig, war alles o.K. - nur leider war unser Tutor nicht wirklich in der Lage, Aufgaben zu erklären.“ „Überflüssig, da Grundlagen in der Theorievorlesung gegeben wurden und die reine Aufgabenbesprechung wenig produktiv war.“

fehlende  
Erklärungen

Einige erkannten, dass das Tutorium von seiner Idee her eine sinnvolle Komponente der Veranstaltung darstellte und Verknüpfungen zu den Vorlesungen herstellte. Auch die Idee des Tutors als Moderator wurde positiv bewertet, hat aber nicht funktioniert.

Als positiv empfanden es viele, nicht alles vorgesetzt zu bekommen wie in anderen Tutorien. Auch die Zusammenarbeit (durch Gruppenbildung) mit den Kommilitonen und die Diskussion im Tutorium sowie die Aufgaben wurden gelobt. „Die Aufgabenblätter waren ein gutes Mittel, um auszuprobieren, ob das Gelernte verstanden und angewendet werden konnte.“

Viele Studierende bemerkten jedoch, dass keine angemessene Nachbereitung der Aufgaben durch die Tutoren stattfand. „Die gröbere Betrachtung der Probleme der Statistik fand nie statt, oft hätte ich mir eine Ergänzung zu der Theorievorlesung gewünscht.“ Auch die Anwesenheitspflicht wurde scharf kritisiert. „... Da die Tutorien oft sehr spartanisch waren und die Korrektur der Aufgaben nur mit geringem Engagement durchgeführt wurde, habe ich mich schon mal gefragt, ob die Teilnahme nicht überflüssig ist.“

Teilnahme  
überflüssig?

**Kommentare zum Reportverfassen:** Das Erstellen statistischer Reports wurde von den Studierenden als sehr sinnvolles Instrument empfunden, um sich statistische Methoden anzueignen. Viele Studierende sahen es als gute Übung an, empfanden es aber als stressig und nervenaufreibend und waren besonders mit der Bewertung sehr unzufrieden - „zeitintensiv, trotzdem weiß ich immer noch nicht, wie ein guter Report geschrieben wird“.

sinnvoll!

Viele Studierende hätten sich einen Report als Musterbeispiel gewünscht. Immer wieder wurde Kritik am unklaren Bewertungsschema geübt und eine Offenlegung der angestrebten Ziele gewünscht. Die Studierenden, die angaben, hier am meisten gelernt zu haben, gaben als Grund an, dass sie während der Reporterstellung merkten, was sie noch nicht richtig verstanden hatten, und bereits Verstandenes vertiefen konnten.

Beispiel  
gewünscht

„Überall ein wenig macht ein Ganzes!“ - fasste ein Student zusammen.

#### 4. Halten Sie die vorgenommene Unterteilung der Vorlesung in Motivation und Theorie für sinnvoll? Begründung

Einige Studierende merkten an, dass ihnen durch die Trennung der Vorlesungen an manchen Stellen der Bezug fehlte. Sie hätten sich eine engere Verbindung der Inhalte gewünscht, vielleicht sogar in einer Vorlesung zusammen. „Durch die Trennung kann man sich zwar besser auf die Motivation und die Theorie im Einzelnen konzentrieren, doch manchmal fehlte durch die Trennung der Bezug zueinander.“ Der Großteil der Studierenden empfand die Trennung als sehr sinnvoll, begründet damit, dass sie sich besser auf die Motivation einlassen konnten, eine Art Überblick bekamen und den Zusammenhang zwischen Statistik und dem Leben sahen und die Theorie expliziter erklärt bekamen. Doch gab es auch vereinzelt gegensätzliche Meinungen. „Gerade bei komplexen Themegebieten halte ich es für sinnvoller, der Motivation die Theorie gleich folgen zu lassen, damit die Motivation nicht verfliegt...“ Einige Studierende merkten an, dass die Trennung der Vorlesung bei ihnen dazu führte, nur die Theorievorlesung zu besuchen. „... Das Problem hierbei ist wohl, dass man den Stoff auch bewältigen kann, wenn man die Motivationsvorlesung auslässt.“ Aus vielen Antworten ließ sich der Bedarf nach einer gewissen Ordnung herauslesen. „Die Unterteilung ist sinnvoll, da man somit einen Plan hat, wie alles läuft, d.h. man hat eine festgelegte Route, der man folgt, um ans Ziel zu kommen.“ Solche Antworten machen deutlich, wie wichtig es ist, den Studierenden zu verdeutlichen, wie die einzelnen statistischen Methoden zusammenhängen, und ihnen einen Überblick zu verschaffen. Erstsemester wollen an die Hand genommen werden.

Bezug  
fehlte

Ordnung  
gewünscht



### 5. Würden Sie sich wiederholt für ein Multimediaprojekt anmelden, wenn Sie die Möglichkeit dazu hätten? Begründung

Alle Studierenden würden sich sofort wieder für ein Multimediaprojekt anmelden und baten um die Weiterführung des Projektes. Sie betonten, viel intensiver als in anderen Fächern gelernt zu haben, führten dies sowohl auf die Anwendung des Multimediaprogrammes als auch auf die Gliederung der Veranstaltung zurück. Viele wiesen aber auch darauf hin, dass es besser wäre, wenn das Projekt nicht ganz so zeitintensiv wäre, um andere Fächer nicht vernachlässigen zu müssen. „Im Prinzip schon, wenn es nicht wieder so ein Stress mit den Aufgaben wäre - das war zu viel!“

Folge-  
projekt  
gewünscht

Sie betonten immer wieder, dass sie es als sehr hilfreich empfanden, gleich im ersten Semester an Multimedia und den Rechner herangeführt zu werden. Einigen merkten positiv an, dass das Projekt dazu führte, dass sie sich auch mit anderen Programmen freiwillig auseinandersetzten, um einen besseren Vergleich zu haben. Die Eigeninitiative wurde stärker gefordert als in anderen Fächern. Die Studierenden sahen es als gute Vorbereitung für das spätere Berufsleben an, Probleme durch den Einsatz von Multimedia lösen zu lernen. Viele betonten aber auch, dass vorrangig die Art der Veranstaltung und die kleine Gruppe dazu beigetragen haben, dass sie viel gelernt haben und jederzeit wieder an einem Multimediaprojekt teilnehmen würden. „Ich würde mich sofort wieder anmelden, aber eher wegen der Gliederung der Veranstaltung und der kleinen Gruppe - das Lernen wird dadurch sehr erleichtert.“

Multimedia  
weckt  
Eigen-  
initiative

Immer wieder stellten die Studierenden ungefragt den Vergleich zur parallel angebotenen, nicht multimedialen Veranstaltung an und betonten, deutlich mehr verstanden und gelernt zu haben als ihre Kommilitonen. „Ja, ich denke das Projekt hat zu einem besseren Lernerfolg geführt.“ Dass solche Vergleiche angestellt werden, ist erstaunlich. Sie sollten jedoch nicht überbewertet werden, da es den Studierenden an Urteilsfähigkeit mangeln dürfte.

Viele Studierende betonten, dass sie durch das Arbeiten mit realen Datensätzen (was nur durch den Einsatz des Rechners möglich war) sehr viel Interesse und Spaß am Lernen hatten. „Ich würde mich auf jeden Fall wieder anmelden. Erstens ist es sowieso interessanter, am Computer herumzuspielen als dicke Bücher zu wälzen, und zweitens ist und bleibt die Anwendung von Multimedia ein wichtiger Bestandteil im Leben bzw. in der Wirtschaft von heute. Dementsprechend ist es wichtig, die Handhabung mit diesen Möglichkeiten zu lernen und anzuwenden.“

...+ Inter-  
esse

### 6. Verbesserungsvorschläge und/oder Kritik

- Die Kritik galt zum Großteil den Tutoren. Rufe nach engagierteren und fachlich qualifizierteren Tutoren werden laut. Besonders die Korrektur der Aufgaben durch die Tutoren stand im Hagelfeuer der Kritik: „... wie soll ich aus meinen Fehlern lernen, wenn Falsches als richtig abgehakt wird!“
- Einige Studierende merkten ernsthaft an, dass sie es für sinnvoller halten würden, die Motivationsvorlesung statt am frühen Morgen in der Mittagszeit stattfinden zu lassen, weil es „die optimale Veranstaltung für zwischendurch“ sei.
- Viele Studierende hätten sich eine intensivere Einführung in die Multimediastsoftware zu Beginn des Semesters gewünscht und kritisierten die betreuten Rechnerstunden, in denen sie sich mehr Aktivität der Tutoren gewünscht hätten, um anhand der Aufgaben das Programm erlernen zu können. Die Tutoren gaben aber keinerlei Hilfestellung bei der Lösung der Aufgaben und waren technisch häufig überfragt oder erklärten so schnell, dass die Studierenden Mühe hatten, ihnen zu folgen. „Nach der Ankündigung von der Projektleitung, dass erstmal keine PC-Kenntnisse erwartet werden und jeder einsteigen kann, weil alles Schritt für Schritt gezeigt wird, hatte ich eine andere, viel gründlichere Einführung in das Programm erwartet.“
- Es wurde der Wunsch geäußert, doch mehr mit dem Programm zu machen, d.h. zum Beispiel nicht nur auf Teile der Theorie zu verweisen und diese anzusprechen, sondern wirklich zusammen in der Vorlesung durchzuarbeiten. Einige Studierende zeigten sich auch in Gesprächen überrascht, wie gut der Theorieteil und das Glossar sich doch

aktivere  
Tutoren!

Einführung  
intensiver

eignen, um sich vor - bzw. nachzubereiten -, sie müssen anscheinend mit der Nase direkt darauf gestoßen werden.

- Viele Studierende waren interessiert an einer Offenlegung des Bewertungsschemas für die statistischen Reports, da sie mit der Beurteilung unzufrieden sind.
- Viele wünschten sich die Vereinigung von Theorievorlesung und Tutorium, da sie mit den Tutoren sehr unzufrieden waren und das Gefühl hatten, in den Tutorien nichts gelernt zu haben. Es ist interessant, dass immer wieder eine klarere Aufgabenstellung bei den Reports gewünscht wurde, was zeigt, dass die Studierenden nicht viel davon halten, ohne klare Anweisung zu arbeiten.
- Auch merkten die Studierenden an, dass es nicht sehr hilfreich sei, wenn im Tutorium der Rechner nicht funktioniert. Dies zeigt, dass organisatorisch in den Hochschulrechenzentren noch eine Menge getan werden muss, bevor flächendeckende Einsätze von Multimedia auch nur im Ansatz möglich sein können. „Der Computer in der Vorlesung und im Tutorium war ein Debakel - und das in einer Multimedieveranstaltung!“
- Ein weiterer Kritikpunkt betrifft die zeitliche Einteilung. Viele Studierenden plädierten dafür, dass 30 Minuten der Motivationsvorlesung an die Theorievorlesung übergeben werden sollten.
- Rundherum positiv wurde die Arbeit in Zweiergruppen hervorgehoben, auch wenn zugegeben wurde, sich manchmal auf den anderen zu verlassen. Viele Probleme wurden in Diskussionen gelöst.
- Einige Teilnehmer würden ein Skript für sinnvoll halten, um gezielter nacharbeiten zu können. „Ihre Kollegen sollten bei Ihnen abgucken und die Theorievorlesungen der Ihrigen entsprechend aufbauen.“

klare  
Aufgaben-  
stellung

Gruppen-  
arbeit

#### 7. Hat Ihnen das Projekt (Reporterstellung) das wissenschaftliche Arbeiten näher gebracht?

Größtenteils sahen die Antworten wie folgt aus: „Durch das Reportverfassen habe ich gelernt, wie bei einem Problem vorgegangen werden muss, um es unter wissenschaftlichen Aspekten zu analysieren und zu lösen.“ „Auf jeden Fall - man lernt die Bibliothek und das Internet schätzen!“

Viele Studierende äußerten, dass es gerade die offene Aufgabenstellung der Reports war, die sie dazu bewegt hat, Literatur heranzuziehen, Annahmen zu treffen, Ziele selbst zu formulieren - auch wenn sie immer wieder den Wunsch nach deutlicheren Aufgabenstellungen äußerten.

Viele konnten mit der Frage aber auch nichts anfangen.

„Es ist schon ein Unterschied, ob man konkrete Aufgaben gestellt bekommt oder sich die Lösung selbst erarbeiten muss. Dies war zum Üben für das wissenschaftliche Arbeiten auf jeden Fall gut.“ „Gegenüber den Aufgabenzetteln musste man beweisen, dass man versteht, worum es geht, welche Probleme zu lösen sind. Als Abstraktion von der Materie waren die Reports gut geeignet, um grobe Zusammenhänge herzustellen oder zu verstehen!“

Viele, die mit der Frage nicht recht etwas anfangen konnten meinten aber, dass es ihnen auf jeden Fall das selbstständige Arbeiten näher gebracht hat.

An dieser Stelle wurde häufig der Wunsch nach einem konkreten Beispiel für einen Musterreport laut.

Selbst-  
ständigkeit

#### 8. Beurteilen Sie die einzelnen Komponenten der eingesetzten Multimediasoftware!

Die Videogeschichte wurde mit Abstand am negativsten beurteilt. „... langweilig, wenig hilfreich, nie ganz angeguckt, ...“ Die Animationen wurden von den Studierenden erst wahrgenommen, als in der Vorlesung darauf hingewiesen wurde. Vorher hatte sich bis auf einen Studierenden noch keiner überhaupt mit ihnen befasst. Dann allerdings äußerten sich die Studierenden sehr positiv über die Animationen und betonten, dass diese ihnen viele statistische Sachverhalte verdeutlicht und zum besseren Verständnis beigetragen hätten. Das

Labor wurde für die Möglichkeiten beim Erstellen von graphischen Darstellungen sehr gelobt. Viele Studierende erkannten besonders beim Reportverfassen die Vorteile des Labors oder überhaupt des rechnergestützten Arbeitens und die Erleichterung, die mit diesem verbunden ist. Ihnen wurde durch das Labor und den Editor bewusst, dass beide unentbehrlich für statistisches Arbeiten sind, auch wenn sie im Einzelnen kleine Schwierigkeiten mit beiden Komponenten hatten. Das Glossar hingegen wurde sehr gelobt, es wurde vorwiegend als Nachschlagewerk genutzt. „... sehr anschaulich und verständlich aufgebaut und erklärt.“ Der Theorieteil wurde vor allem als Nachschlagewerk statt Literatur und zur Ergänzung der Vorlesungen benutzt. Besonders positiv und hilfreich wurden die Beispiele beurteilt. Einige Studierende kritisierten die unmathematische Gestaltung des Theorieteils, was andere wiederum als Vorteil betrachteten. „... Die Aufmachung ist oft sehr unmathematisch, was manchen Leuten zusagt, aber eigentlich sollte die grundlegende Mathematik zumindest vorhanden sein, da man sie benötigt.“ Ansonsten äußerten sich die Studierenden größtenteils positiv über den Theorieteil, obwohl sie immer erst auf diesen hingewiesen werden mussten und ihn dann nutzten, nicht von allein, wie man vermuten könnte.

Erleichterung durch Labor

Lob ans Glossar

Viele äußerten sich negativ darüber, dass nur die deskriptive Statistik durch die Multimedia-Software abgedeckt wurde. Die im Theorieteil verankerten Aufgaben und Lösungen wurden als gute Möglichkeit angesehen, Kenntnisse zu überprüfen.

Mehrfach wurde kritisch angemerkt, dass sich das Programm immer in den Vordergrund schiebt und durch das Hantieren mit der Taskleiste häufig Abstürze provoziert werden. Die Idee, den Editor als Reportgenerator einzusetzen, wurde gelobt. Allerdings wurden die technischen Schwierigkeiten (Druckprobleme, copy-paste Probleme, ...) kritisiert. Aus einigen Ausführungen ging deutlich hervor, dass die Studierenden fast ausschließlich das Labor benutzten und gelegentlich den Editor, die anderen Komponenten aber nicht wahrnahmen.

Technische Probleme

Die folgenden Äußerungen zeigen, dass die Meinungen der Studierenden bezüglich *Statistik interaktiv* divergieren.

Meinungen

„Recht gelungen erscheint mir der LDL-Kalkulator, der die Möglichkeit gibt, eigene Funktionen zu programmieren oder kleinere Programme zu erstellen.“

„Leider konnte ich mich mit dem Programm nicht recht anfreunden, speziell mit dem Editor stand ich auf Kriegsfuß, vergleicht man das Labor mit MS-Excel, so kann man in Excel sehr viel unkomplizierter Berechnungen durchführen.“

„Gute Ansätze erkennbar, aber viele fehlerbehaftete Module, instabiles Laufverhalten, ...“

„... Das Programm ist zu komplex, um den Umgang nebenher zu erlernen.“

„Animationen: immer noch nicht angeschaut - mach ich aber noch!“ „Die Animationen erinnern an die Sendung mit der Maus. Mich nervt es, wenn man mir etwas erklärt, als wäre ich noch ein Kleinkind!“

„... Die Software war sehr hilfreich beim Bewältigen der Aufgaben, eine willkommene Hilfe. Vor allem das Labor war ein sehr einfach zu bedienendes und hilfreiches Mittel, aber auch das Glossar war ein äußerst nützliches und ausführliches Nachschlagewerk.“

**9. Welche der in der Vorlesung vermittelten statistischen Gebiete finden Sie am interessantesten, z.B. weil Sie sich vorstellen können, sie später noch gebrauchen zu können?**

Ein Großteil der Studierenden äußerte generelles Interesse an allen Gebieten und wies darauf hin, dass verstanden wurde, wie oft die Statistik eine Rolle spielt, und dass statistische Kenntnisse von Nutzen sind. Oft wurde auch angemerkt, dass große Schwierigkeiten bezüglich des Verstehens der Wahrscheinlichkeitsrechnung bestanden, trotz ausführlicher und vielseitiger Behandlung in den Veranstaltungen. Es wurde aber meist im gleichen Satz erwähnt, wie interessant diese sei, da sie zu häufig verblüffenden Ergebnissen geführt habe. Es wurde den Studierenden bewusst, dass es wichtig ist, sich mit den Grundlagen der Wahrscheinlichkeitsrechnung auszukennen, um Situationen nicht intuitiv falsch einzuschätzen.

WS-rechnung

Viele Studierende nutzten die Beantwortung dieser Frage zu einem generellen Fazit. Sie lobten den ständigen Praxisbezug und das Arbeiten mit realen Datensätzen. Sie fühlten sich

durch diese Arbeit sehr zum Lernen motiviert. Den Zusammenhang verschiedener Merkmale zu untersuchen oder sich mit Hilfe der 5-Zahlen-Zusammenfassung/Boxplot einen ersten Überblick über die Daten zu verschaffen, wurde als interessant hervorgehoben. Auch der Normalverteilung wurde besonders bezüglich ihrer Rolle in der Statistik mit Interesse begegnet. Immer wieder erwähnten die Studierenden an dieser Stelle das Reporterstellen, das ihnen am meisten gebracht habe, da sie davon überzeugt sind, dies im Berufsleben auch beherrschen zu müssen.

„Als ich mich auf die VWL-Klausur vorbereitete und der Phillipskurve begegnete, dachte ich, wäre ich vor Phillips geboren, hätte ich diese Kurve auch entdecken können. Er hat bloß eine statistische Methode benutzt und ein Streudiagramm gezeichnet.“

## G Qualitative Befragung der 2. Durchführung

In dem zweiten Fragebogen fand die angestrebte Kombination quantitativer und qualitativer Evaluationsfragen statt. Es folgt die Auswertung der qualitativen Evaluationsfragen.

*Konnte Ihr Interesse für Statistik durch das Projekt geweckt werden? Wenn ja, ist dies eher auf den Einsatz von Multimedia zurückzuführen oder auf andere Aspekte wie zum Beispiel die Betreuung, Form der Vorlesung, ...*

Die Antworten der Studierenden zeigen, dass bei einem Großteil das Interesse für Statistik geweckt werden konnte. Dies ist gleichermaßen auf drei Aspekte und deren Kombination zurückzuführen:

- die Gestaltung der Veranstaltung: Vorlesung, Übung, Rechnerstunde in kleiner Gruppe, Arbeit in Teams,
- den Einsatz der multimedialen Lernsoftware *Statistik interaktiv komplett*, insbesondere des Statistiklabors,
- die starke Betreuung der Studierenden in Verbindung mit der Nähe zu den jungen Dozenten.

Viele Studierende betonten, dass sie durch den Einsatz der multimedialen Lernmodule motivierter an die statistischen Inhalte herangegangen sind, insbesondere bei der Nacharbeitung der Vorlesungen. Das Leitfadenprinzip wurde in diesem Zusammenhang gelobt, da es eine Organisation der verwendeten/verwendbaren<sup>323</sup> Module bereits vornahm. Ein Großteil der Studierenden bezeichnete das Arbeiten mit verschiedenen Modulen/Komponenten als interessant und abwechslungsreich, gab aber auch zu, Startschwierigkeiten gehabt zu haben. „Der Einsatz von vielen Medien war am Anfang etwas verwirrend. Man brauchte einige Zeit, sich zurechtzufinden.“ Viele betonten, dass besonders das Arbeiten im Labor, in der Vorlesung und zu Hause ihr Interesse für die Statistik geweckt hat, sie das Statistiklabor als ein Instrument zur Arbeitserleichterung empfunden haben. Allerdings ist aus den Antworten deutlich herauszuhören, dass der Hauptgrund für ihr Interesse an der Statistik in der Betreuung liegt. „Die Atmosphäre in den Vorlesungen war sehr angenehm und die Dozenten haben super erklärt.“ „... lockere Form der Vorlesung, Arbeiten im Team“. Die angenehme Atmosphäre wurde insofern gelobt, als man sich nicht scheute, Fragen zu stellen. Die kleine Gruppe, die eine starke Betreuung des Einzelnen ermöglichte, wurde positiv hervorgehoben. Zusammenfassend muss festgestellt werden, dass die multimedialen Lernmodule zwar einen positiven Effekt auf das Interesse der Studierenden für die Statistik hatten, dass aber die starke Betreuung durch junge Dozenten deutlich höher bewertet wurde.

*Empfanden Sie es als positiv, Teile der Prüfungsleistung schon während des Semesters erbringen zu können (Aufgaben, Reports) - wenn ja, warum? ... Oder empfanden Sie die höhere Belastung während des Semesters als eher störend (Gründe)?*

Ein Großteil der Studierenden nahm eine positive Bewertung vor. Es wurde betont, dass die Reports während des Semesters eine kontinuierliche Auseinandersetzung mit dem Stoff hervorgerufen haben, was aber auch sehr arbeitsaufwendig war. „Man ist dazu gezwungen, sich gleich nach der

<sup>323</sup>Einige Module wurden in der Vorlesung empfohlen, dort aber nicht eingesetzt (z.B. Animationen).

Vorlesung mit dem Stoff zu beschäftigen, was man sonst nicht tut.“ „Die Belastung war auf jeden Fall höher als in anderen Fächern. Allerdings hat man dadurch, dass man sich durchgängig und intensiv mit dem Thema befasst hat, auch mehr gelernt.“ Die Studierenden berichteten zwar über die als stark empfundene Belastung während des Semesters, hoben aber im selben Atemzug positiv hervor, dass sie sich zu Semesterende besser auf die restlichen Grundstudiumsklausuren konzentrieren konnten. Die Abgabe der wöchentlichen Übungsaufgaben wurde in diesem Zusammenhang ähnlich bewertet. Die Belastung war groß, aber man setzte sich kontinuierlich mit dem Stoff auseinander. Interessant ist die Beobachtung, dass sich die ausländischen Teilnehmer durchweg positiv äußerten und nicht über eine zu hohe Belastung klagten, sondern das kontinuierliche Lernen hervorhoben. Kritisch sollte allerdings die Anmerkung einiger Teilnehmer betrachtet werden, durch den hohen Arbeitsaufwand in Statistik seien andere Fächer zu kurz gekommen. Einige Teilnehmer schlugen vor, den Abschlussreport einige Wochen früher, also vor Semesterende zu schreiben, um ihn aus der Vorbereitungszeit für die anderen Klausuren herauszunehmen. Ein weiterer Vorschlag bestand darin, die wöchentlich abzugebenden Lösungen der Übungsaufgaben stärker gewichtet in die Endnote einzubeziehen, da diese einen großen Teil der Arbeit ausmachten.

*Das Projekt besteht aus verschiedenen Komponenten: Vorlesung, Übung, Rechnerstunden, selbstständige Reporterstellung - welche Meinung haben Sie bezüglich der einzelnen Komponenten (Stichworte), wo haben Sie am meisten gelernt? Verbesserungsvorschläge/Kritik?*

Die Studierenden betonten, dass eigentlich alle Komponenten wichtig waren und zusammen positiv gewirkt haben. Viele vertraten die Meinung, dass das selbstständige Verfassen der Reports einen hohen Beitrag zum Lernerfolg geleistet habe. Auch ein Großteil bewertete die in den letzten 20 Minuten jeder Vorlesung gemeinsam gelösten Beispielaufgaben zum vorher vermittelten Stoff als sehr hilfreich. Einige Studierende bewerteten alle Komponenten in Verbindung miteinander und betonten, dass das Prinzip Vorlesung-Übung-Rechnerstunde-Reporterstellung sehr hohe Lernerfolge hervorgerufen habe. Sie führten das zurück auf das Zusammenwirken und die zeitliche Abfolge (langsame Vertiefung, schrittweises selbständiges Arbeiten) der einzelnen Komponenten. Viele betonten, dass besonders die Unterstützung in der Rechnerstunde dazu beigetragen habe, dass sie den Umgang mit dem Statistiklabor erlernten und zur Lösung der statistischen Probleme einsetzen konnten. Positiv wurde außerdem hervorgehoben, dass die Aufgaben in der Vorlesung und Übung gemeinsam mit dem Dozenten erarbeitet wurden, also im Dialog gelernt wurde. Einige Studierende hätten in der Übung gern mehr im Statistiklabor gearbeitet. Es bleibt zu überlegen, auch die Übung in einem Rechnerpoolraum abzuhalten, um Aktivität der Studierenden am Rechner zu ermöglichen. Kritisch wurde die Fülle des zu vermittelnden und zu erlernenden Stoffes angesprochen. Die Studierenden fühlten sich durch die wöchentlich neu vermittelten statistischen Inhalte/Methoden überfordert.

Zusammenfassend kann festgestellt werden, dass die Rollen der einzelnen Komponenten erkannt und positiv bewertet wurden. Wünschenswert hätten es einige gefunden, ein größeres Angebot an betreuten Rechnerstunden angeboten bekommen zu haben.

*Würden Sie sich wiederholt für ein Multimediaprojekt anmelden, wenn Sie die Möglichkeit dazu hätten? (Begründung)*

Den Antworten ist zu entnehmen, dass ein Großteil der Studierenden jederzeit wieder an einem Multimediaprojekt teilnehmen würde. Es wurde häufig nach einem Anschlusskurs *Statistik interaktiv II* gefragt. Eine häufig genannte Begründung war das interessantere Lernen durch eigene Aktivität in den Komponenten der Lernmodule. Ein weiterer Grund war die Arbeit in kleinen Gruppen, die von den Studierenden in Verbindung gebracht wurde mit einem Multimediaprojekt. „... angenehmes Lernen durch gute Betreuung, man fühlt sich nicht so sehr allein mit Problemen wie in vollen Vorlesungen.“ „Ja, weil es schön war, auch mal in einer kleinen Gruppe zu lernen und weil man viel mehr Kontakt zum Dozenten hatte.“

Außerdem bezeichneten viele das Arbeiten in einem Multimediaprojekt als interessanter und abwechslungsreicher im Vergleich zu den anderen Vorlesungen. Einige betonten, großes Interesse an Multimediaprojekten zu haben, da sie so besser auf die berufliche Zukunft und den damit ver-

bundenen Umgang mit neuen Medien vorbereitet werden. Ein interessanter Beitrag lautet: „Ja, da der Umgang mit Medien immer wichtiger wird, Bearbeitung ist viel einfacher (Grafiken etc.) und spart viel Zeit, die für wichtigere Erklärungen verwendet werden können.“ Genau dieses Ziel wurde mit dem Einsatzkonzept verfolgt. Es scheint sich erkennbar durchgesetzt zu haben. Die Prüfung in Form der statistischen Reports wurde von vielen als weiterer Grund genannt, sich wieder für ein solches Projekt anzumelden. Außerdem betonten viele, deutlich mehr als in anderen Veranstaltungen gelernt zu haben, führten dieses auf die Veranstaltungsform (Multimediaprojekt) zurück und würden sich deshalb jederzeit wieder für ein solches Projekt anmelden.

*Hat Ihnen das Erstellen statistischer Reports das wissenschaftliche Arbeiten näher gebracht?*

Die Antworten lassen darauf schließen, dass die Studierenden mit dem Begriff „wissenschaftliches Arbeiten“ Schwierigkeiten haben. Einige betonten, dass sie gelernt haben, ihre Lösungen zu strukturieren, Methoden ausgewählt und begründet anzuwenden. Außerdem bemerkten einige, dass sie gelernt hätten, ein Problem zu beschreiben, einzugrenzen und Analysen in Bezug auf das Problem zu interpretieren. Sie betonten, dass dies in keinem anderen Fach im Grundstudium vermittelt werde, da es dort hauptsächlich um die Lösung von Übungsaufgaben, die Anwendung von Kochrezepten gehe. Ein Teil der Antworten lässt vermuten, dass die Studierenden wissenschaftliches Arbeiten gleichsetzen mit statistischem Problemlösen, d.h. dass sie nicht von der Statistik abstrahieren. Der Aspekt der Dokumentation des eigenen Arbeitens wurde von vielen als wichtig in Bezug auf wissenschaftliches Arbeiten betrachtet. „Ich achte jetzt viel genauer darauf, was ich eigentlich schreibe.“ Dies kann zurückgeführt werden (und wird von den Studierenden zurückgeführt) auf die detaillierten Reportbesprechungen, in denen einzelne Sätze der Studierenden kritisch diskutiert wurden. Einige erklärten, sich im Nachhinein in Vorbereitung auf die mündliche Prüfung kritisch mit ihren eigenen Texten (Reports) auseinandergesetzt zu haben. Es wurde angemerkt, dass das wissenschaftliche Arbeiten im Laufe des Semesters durch die Reportbesprechungen, die Anfertigung der Minireports, der Reports schrittweise verbessert wurde. Die Idee der Verankerung eines Beispielreports in den Lernmodulen scheint positive Effekte auf die Reporterstellung der Studierenden zu haben. Einige bemerkten, dass ihnen dieser eine große Hilfe war.

*Beurteilen Sie die einzelnen Komponenten der eingesetzten Multimediastsoftware!*

Die Beurteilung erfolgte komponentenweise, wie die folgende Dokumentation der Antworten zeigt:

- Glossar: technische Probleme
- Beispiele: gut für Anwendungsverständnis
- Applets: gute Anschauungskraft
- Statistiklabor: Technische Mängel werden angesprochen, aber ansonsten wurde das Statistikkabor als eine sehr wichtige Komponente des statistischen Problemlösens betrachtet. Ein Studierender machte den Vorschlag, eine Formelsammlung für den R-Kalkulator zu entwickeln und zu implementieren. Obwohl an dieser Stelle nach den Multimediakomponenten gefragt wurde, bemerkten viele Studierende, dass ihnen die Vorlesungen und Übungen am meisten gebracht hätten. Der Umgang mit dem R-Kalkulator wurde von vielen als schwierig bezeichnet. Sie betonten, mehr Hilfestellungen zu benötigen, um mit dem R-Code hantieren zu können. Die Studierenden sahen, dass sich das Labor noch in keiner Endversion befindet und betonten, dass es das wichtigste Element der Lernsoftware sein werde, sobald alle technischen Mängel beseitigt seien. Sie sprachen ganz konkrete Mängel an, z.B. dass die Konnektoren nicht über die Bildschirmseite hinaus eingesetzt werden können (scrollen). Der R-Grafik-Wizard wurde gegenüber dem R-Kalkulator positiv bewertet, der Umgang sei angenehmer für die Studierenden, der R-Kalkulator wurde nicht als Instrument mit mehr Möglichkeiten erkannt. Der generelle Aufbau des Statistikkabors wurde positiv hervorgehoben (Objekte, Klicken, ...).
- Theoriekomponenten eingesetzt zur Nachbereitung der Vorlesungen, durchgehend als hilfreich empfunden.

- Blackboard: Die gute Organisation des Lernens wurde von vielen Studierenden hervorgehoben, besonders die (Online-)Kommunikation, die viele als sehr zeitgemäß empfinden.

Es ist auffallend, dass die Studierenden keine Unterscheidung zwischen dem Lernmanagementsystem Blackboard und den Theoriekomponenten der Lernsoftware *Statistik interaktiv komplett* vornahmen. Aus den Antworten wird deutlich, dass sie die Theoriekomponenten als Teile des Lernmanagementsystems betrachteten, nicht sahen, dass die Lernsoftware in dieses eingebettet ist. Blackboard wurde neben diesem Aspekt zusätzlich sehr positiv als Kommunikationsplattform zwischen Studierenden und Dozenten hervorgehoben.

Die allgemeinen Bemerkungen der Studierenden sind nicht einheitlich. Einige bezeichneten die Lernmodule als übersichtlich, hoben die Vernetzung der Komponenten durch Hyperlinks positiv hervor, andere kritisierten gerade diese Aspekte, schlugen vor, sich auf das Wesentliche zu beschränken, sprachen sich gegen die Hyperlinks aus. Einige betonten, dass die Lernmodule gerade im Verbund sinnvoll seien, generell ein Selbststudium ermöglichen und anschaulicher sind als Literaturquellen, auf die in der Vorlesung auch verwiesen wurde.

*Welche der in der Vorlesung vermittelten statistischen Gebiete finden Sie am interessantesten, z.B. weil Sie sich vorstellen können, sie später noch gebrauchen zu können?*

Die Antworten sind sehr unterschiedlich. Ein Teil der Studierenden erachtete alle Methoden in Verbindung miteinander als notwendig, merkte z.B. an, dass die Normalverteilung allein nicht viel Erkenntnis bringe, wenn man sich nicht mit Hilfe deskriptiver Methoden einen Überblick über die Daten verschaffen kann. Andere griffen einzelne Gebiete heraus und begründeten, warum sie gerade diese als interessant und später relevant (Berufsleben) erachteten:

- Wahrscheinlichkeitsrechnung - enger Bezug zum alltäglichen Leben, außerdem in den anderen Gebieten benötigt
- Grafische Darstellungen, deskriptive Statistik - Überblick über Daten verschaffen/gewinnen, erste Analysen durchführen
- Bivariate Datenanalyse - Zusammenhänge zwischen Merkmalen untersuchen

Die Wahrscheinlichkeitsrechnung wurde von einem Großteil der Studierenden genannt, wenn auch häufig ohne Begründung.

*Wie schätzen Sie das Teamwork ein? (Begründung)*

Ein Großteil der Antworten spiegelt eine sehr positive Bewertung des Teamworks wider. „Ist sehr sehr wichtig und auch gut. man sollte lernen, im Team zu arbeiten, weil sich zwei Leute immer sehr gut ergänzen.“ Der Aspekt des gegenseitigen Ergänzens, insbesondere in Bezug auf die Erstellung der ersten beiden Reports taucht in vielen Antworten auf. „Gut, es ist auch für später wichtig, im Team arbeiten zu können. Man kann sich gegenseitig helfen und Probleme diskutieren. Die Arbeitsergebnisse sind besser.“ Wenige Studierende äußerten Kritik an ihrem Partner, fühlten sich ausgenutzt, hätten die Arbeit allein gemacht. Generell ist die Kritik allerdings äußerst positiv.

In diesem Zusammenhang wird kritisiert, dass das gemeinsame Verfassen statistischer Reports das ganze Semester über trainiert werde und dann im Abschlussreport nicht konsequent weitergeführt werde. Die Studierenden betonten, im Team sehr viel produktiver arbeiten zu können, da man sich gegenseitig diszipliniert, sich zuvor mit dem Stoff auseinanderzusetzen. Es wird deutlich, dass ein Großteil der Teams sehr eng zusammengearbeitet hat, das Ziel erreicht wurde, Diskussionen und kritische Reflexionen anzustoßen. Die Studierenden berichteten über das gegenseitige Aufdecken und Beseitigen von Fehlern und Verständnisschwierigkeiten, Sammeln von Ideen zur geeigneten Lösung statistischer Problemstellungen. Interessant sind auch Bemerkungen dahingehend, dass nicht nur innerhalb der gebildeten Zweierteam, sondern auch zwischen verschiedenen Teams Kommunikation stattfand. Dies sollte man kritisch betrachten, da die statistischen Reports streng genommen eine Prüfungsleistung darstellen.

*Sehen Sie die Erstellung der statistischen Reports als gute Einführung in das statistische Problemlösen? (Begründung)*

Die Antworten der Studierenden sind durchgehend positiv. Sie bemerkten, dass sie durch die Anfertigung der Reports gelernt haben, wie sie an das Lösen statistischer Probleme herangehen sollten. „Ja, man befasst sich mit einem Problem, entscheidet über Methoden, dieses zu lösen, und kommt schließlich zu einem Lösungsvorschlag.“ „Ja, denn es ist nicht einfach, Daten systematisch zu analysieren und erfordert viel Übung.“ „Ja, da man sich so umfangreich Gedanken um ein Problem macht und dieses auch für den Alltag sehr interessant sein kann.“

Außerdem erlernten sie, strukturierte Lösungen zu erstellen, Analysen zu interpretieren und ihr Arbeiten zu dokumentieren. Durch die Reports verstanden die Studierenden nach eigenen Aussagen die Auswahl geeigneter Methoden und merkten, dass nicht immer alle erlernten Methoden zielgerichtet angewendet werden können. Durch die problembezogene Methodenanwendung diskutierten sie verschiedene Möglichkeiten der einzelnen Methoden und bemerkten die Notwendigkeit einer begründeten Auswahl. Die Studierenden betonten, dass es sehr hilfreich war, in der Coachingkomponente *Statistisches Reportverfassen* die Struktur eines statistischen Reports, die einzelnen Phasen und ihre Ziele und wichtigsten Aspekte nachlesen zu können. Erkannt wurde, dass es auch zum statistischen Problemlösen gehört, angemessen zu argumentieren. Dies zeigen Aussagen der folgenden Art: „Ja, man lernt, seine Methoden besonnen auszuwählen, auf die Wortwahl zu achten sowie den Sinn einer Untersuchung und der Daten zu hinterfragen.“ Einzelne erkannten, dass sie die erlernte Vorgehensweise beim statistischen Problemlösen auf das Lösen aller Probleme übertragen können, die erlernten Fähigkeiten sowohl im Studium als auch im späteren Berufsleben anwenden können. Ein Student trifft den Nagel auf den Kopf: „Zum Erstellen eines statistischen Reports muss man verstanden haben, was man tut, und es begründen. Das hilft beim Verstehen.“ Andere betonten, dass nicht nur die Erstellung der statistischen Reports, sondern insbesondere die kritische Nachbesprechung zum Erlernen des statistischen Problemlösens beigetragen haben. Eine Bemerkung ist in fast jeder Antwort enthalten. Es hat die Studierenden motiviert, reale Probleme mit Hilfe statistischer Methoden zu lösen, d.h. zunächst die Theorie zu verstehen, um sie dann praktisch anzuwenden.

## H Multiplechoiceklausur

Kreuzen Sie von den folgenden Aussagen die Ihrer Meinung nach richtigen Aussagen an.

- Das Maximum von Beobachtungen hat einen großen Einfluss auf den Interquartilsabstand.
- In der Regel gilt: Mit wachsender Spannweite wächst auch die Stichprobenvarianz.
- Die Stichprobenstandardabweichung  $s$  ist in keinem Fall negativ.
- Für die Symmetrie-Frage ist die Betrachtung der Differenz aus Mittelwert und Median nicht hilfreich.
- $s^2$  sei die Stichprobenvarianz und  $d^2$  die mittlere quadratische Abweichung, dann gilt:

$$s^2 = d^2 + \frac{n}{n-1} (\overline{x^2} - \bar{x}^2)$$



Welche der folgenden Maße sind Streuungsmaße?

- Median
- Stichprobenvarianz
- Spannweite
- mittlere quadratische Abweichung
- getrimmtes arithmetisches Mittel

Ein Boxplot ist eine graphische Darstellung, die Auskunft gibt über:

- relative Häufigkeiten
- kumulierte Häufigkeiten
- Quartile, Minimum und Maximum
- die Spannweite
- Beobachtungsränge

Hinweis: Angekreuzte falsche Alternativen führen zu einem Punkteabzug.

Kreuzen Sie von den folgenden Aussagen die Ihrer Meinung nach richtigen Aussagen an.

- Die Summe von 10 unabhängigen Bernoulli-verteilten Zufallsvariablen mit dem Parameter  $p = 0.1$  ist binomialverteilt mit den Parametern  $n = 10$  und  $p = 10\%$ .
- Der Erwartungswert einer Zufallsvariablen mit der Dichtefunktion  $f(x) = \lambda e^{-\lambda x}$  mit  $x \geq 0$  ist  $\lambda$ .
- Die geometrische Verteilung ist das diskrete Gegenstück zur Exponentialverteilung.
- Die Poisson-Verteilung kann zur Approximation der geometrischen Verteilung verwendet werden.
- Die Anzahl der in einer Stunde bei einer Bank abgegebenen 1-DM-Stücke sei Poisson-verteilt mit dem Parameter  $\lambda$ . Dann ist die Anzahl der in 20 Minuten abgegebenen 1-DM-Stücke Poisson-verteilt mit dem Parameter  $\lambda/3$ .

Betrachten Sie zum Abschluß eine Spardose mit 50 unterschiedlichen Münzen aus den verschiedensten Ländern. Es werden 5 Münzen zufällig und ohne Zurücklegen entnommen. Die Anzahl möglicher Stichproben beträgt:

- $50^5$
- $\binom{50}{5}$
- $\frac{50!}{50 \cdot (50-5)!}$
- $\frac{50!}{5! \cdot 45!}$
- $50 \cdot 49 \cdot \dots \cdot 46$

## I Technische Ausführungen zu Lernmanagementsystemen

In diesem Abschnitt befinden sich Ausführungen zu technischen Problemen im Einsatz. Es wird zuerst die Ausgangssituation betrachtet. Hier werden die eingesetzten Techniken aufgeführt. Danach werden für die jeweiligen Techniken Ausfallszenarien mit ihren Auswirkungen beschrieben. Zuletzt wird ein Versuch unternommen, Schlussfolgerungen und Lösungen aus den Ausfallszenarien zu ermitteln.

### Ausgangssituation

Die Veranstaltung besteht aus verschiedenen Komponenten: der Vorlesung, der Übung und der Bearbeitung der Inhalte durch die Studierenden. Die Vorlesung findet im Hörsaal statt, der mit einem Computer mit Netzanschluss und einem Projektor ausgestattet ist. Die Übung wird in einem Rechnerpoolraum durchgeführt. Dieser enthält mehrere Computer mit Netzanschluss. Der Studierende bearbeitet seine Aufgaben entweder in diesem Poolraum oder an seinem heimischen Computer, der über einen Netzanschluss verfügt.

Für die Veranstaltung wird ein Lernmanagementsystem eingesetzt. Es übernimmt die Aufgaben, die Veranstaltungsmaterialien an die Studierenden zu verteilen, die Kommunikation zwischen den Studierenden und dem Dozenten zu ermöglichen, die Übermittlung von Dokumenten zwischen den Studierenden und Dozenten zu verwalten. Sämtlicher Austausch von Informationen zwischen dem Dozenten und den Studierenden wird über das Lernmanagementsystem abgewickelt.

Als zweite Komponente nimmt das Statistiklabor eine zentrale Rolle in der Veranstaltung ein. Es wird sowohl durch den Dozenten in der Vorlesung / Übung wie durch den Studierenden bei der Nachbereitung der Veranstaltung und der Bearbeitung von Übungsaufgaben verwendet.

Für eine Veranstaltung mit dem Lernmanagementsystem *Blackboard* und dem Statistiklabor sind folgende technischen Voraussetzungen notwendig: Das LMS läuft auf einem Server, der über Netzwerk zu erreichen ist. Folglich findet die gesamte Kommunikation über ein Netzwerk und das Internet statt. Der LMS-Server muss deshalb nicht am Veranstaltungsort stehen. In unserem Fall befindet sich der Server an der FU Berlin. Das Statistiklabor muss auf den Rechnern der Dozenten und Studierenden lokal installiert sein. Dieser Clientrechner nimmt den Kontakt mit dem LMS-Server mittels eines Browsers auf.

Da sämtliche Materialien nur auf dem LMS-Server zur Verfügung stehen, werden während einer Vorlesung oder Übung die benötigten Informationen von dem Server geladen.

Die Veranstaltung besteht aus Vorlesungen und Übungen, die regelmäßig zu bestimmten Terminen stattfinden. Die Durchführung dieser Termine sind jedoch nur in einer funktionierenden Umgebung möglich.

Es kann nicht sein, dass ein Termin ausfallen muss, nur weil eine der Techniken nicht einwandfrei arbeitet.

### Ausfallszenarien

Bei dem Einsatz der verschiedenen Techniken kann es zu unterschiedlichen Ausfällen kommen, die mehr oder weniger gravierende Auswirkungen auf die Funktionsfähigkeit der Lehrveranstaltung haben können. Dabei können Probleme in der Netzinfrastruktur und den beteiligten Komponenten entstehen:

- **Netzausfall:** Symptom für diesen Fall ist, dass kein Server im Netz erreicht werden kann. Dafür kann es verschiedene Erklärungen geben. Das lokale, hier universitätsinterne Netzwerk ist ausgefallen, oder die Netzverbindung über das Internet zu dem LMS-Server kann nicht aufgebaut werden.
- **Serverausfall:** Die Netze funktionieren, jedoch kann der LMS-Server nicht erreicht werden. Er ist aus welchen Gründen auch immer außer Betrieb.
- **lokaler Rechnerausfall:** Der Rechner z.B. im Hörsaal ist ausgefallen. Es ist keine Beamerpräsentation möglich.

### Konsequenzen / Lösungen

Die Konsequenz eines jeden Ausfallszenarios besteht darin, dass die Veranstaltung nicht wie geplant durchgeführt werden kann, weil keine Informationen mit dem LMS-Server ausgetauscht werden können. Damit ist ein großes Problem durch die zentrale Datenhaltung und die damit verbundene Abhängigkeit von der Möglichkeit, mit dem zentralen Server Verbindung aufnehmen zu können, entstanden.

Zu den einzelnen Szenarien:

- Der lokale Netzausfall unterbricht sowohl bei einer Unterbrechung der Internetverbindung als auch bei einem lokalen Netzwerkausfall die Verbindung zu jedem anderen Rechner. Folglich können keine Daten, die nicht auf dem lokalen Computer zur Verfügung stehen, verwendet werden.

Das hat in den einzelnen Komponenten der Veranstaltung Auswirkungen. Für die Vorlesung im Hörsaal bedeutet es, dass man den Rechner nur lokal benutzen kann, sofern man sich wegen einer zentralen Benutzerverwaltung überhaupt an dem Rechner anmelden kann. Durch die Verwendung durch viele Dozenten ist die lokale Speicherung von Daten nur schwer möglich und kaum zu verwalten. Somit fällt hier die Möglichkeit einer kompletten lokalen Vorhaltung der notwendigen Daten aus. Der Dozent kann nur die notwendigen Materialien für den einen Veranstaltungstermin vorbereiten und dann lokal aufspielen. Hier kann es Probleme mit der Datenmenge geben, so dass diese Möglichkeit nur über eine Speicherung der Information auf CD möglich macht. Eine solche Master-CD mit allen Materialien und der notwendigen Software könnte vor Beginn der Veranstaltung erstellt werden. Sie ermöglicht es dem Dozenten, jedenfalls eine Veranstaltung durchzuführen. Jedoch können nicht aktuellere Materialien und Veränderungen der Datenbestände verwendet werden.

Für die Übungen in den Poolräumen hat es dieselben Auswirkungen. Auch hier können die Anwender nicht an die notwendigen Materialien gelangen. Im Vergleich zu dem Rechner im Hörsaal besteht für den Studierenden hier nicht die Möglichkeit, die Materialien lokal aufzuspielen. Eine Verwendung einer Master-CD für Übungen ist nur soweit sinnvoll, wie auf allgemein zugängliche Materialien der Veranstaltung zugegriffen werden soll. Eine Bearbeitung von Aufgaben durch den Studierenden ist jedoch nicht möglich, weil er nicht auf seine Daten zugreifen kann. Deshalb ist in diesem Fall eine Durchführung einer Übung unmöglich.

Eine häusliche Bearbeitung der Aufgaben durch die Studierenden ist von diesem Szenario nicht betroffen. Das Problem besteht jedoch in den ungleichen Möglichkeiten der Studierenden, die zu Hause über einen Rechner mit Internetzugang verfügen, und den Studierenden, die auf die Arbeitsumgebung der Poolräume angewiesen sind. Dasselbe gilt auch für die Dozenten und Tutoren. Auch sie könnten nur außerhalb der Universität die Veranstaltung vorbereiten und betreuen. Das ist nicht zumutbar.

- Ein Ausfall des Internets verhindert eine Kontaktaufnahme mit dem entfernten LMS-Server. Das hat erstmal die gleichen Konsequenzen wie der Ausfall des lokalen Netzwerks. Jedoch ist jetzt auch das Arbeiten außerhalb der Universität unmöglich geworden. Wenn das interne Universitätsnetzwerk am Veranstaltungsort noch funktionsfähig ist, besteht die Möglichkeit einen lokalen LMS-Server mit redundanten Inhalten zu betreiben. Diese Möglichkeit erfordert jedoch, dass man den entfernten LMS-Server spiegelt.

Für die Vorlesung bedeutet das, dass man auf die Materialien wie gewohnt zugreifen kann. Dieser Teil der Veranstaltung ist demnach nicht betroffen. In der Übung arbeitet man aber mit sich ständig ändernden Datenbeständen, da die Studierenden nach und nach ihre Lösungen erarbeiten. Deshalb erfordert diese Komponente ein regelmäßiges Spiegeln der Inhalte des Lernmanagementsystems. Über die Abstände kann man diskutieren.

Das wirft natürlich die Frage auf, warum der LMS-Server nicht immer am Veranstaltungsort platziert wird. Dagegen sprechen der hohe Administrationsaufwand sowie die Kosten.

Die Materialien der Master-CD können auch über das lokale Netzwerk ohne einen LMS - Server verteilt werden. Dies erfordert jedoch die Einbindung der CD in das lokale Netzwerk oder in einen lokalen Webserver.

- Der Serverausfall verhindert die Verwendung des Lernmanagementsystems. Für diesen Fall ist die folgende Lösung denkbar: Am entfernten Standort des Servers könnten redundante Fallbacklösungen aufgebaut werden, die bei einem möglichen Serverausfall einspringen. Damit kann man am Veranstaltungsort ohne Probleme weiterarbeiten.
- Der lokale Rechnerausfall betrifft nicht die Funktionsfähigkeit des Lernmanagementsystems. Er zeigt nur die Abhängigkeit einer multimedialen, rechnergestützten Veranstaltung vom Medium Computer. In diesem Fall sind weder die Vorlesung, die Übung noch das häusliche Arbeiten möglich. Man muss sich auf die traditionellen Medien zurückziehen. Die Frage ist, ob für diesen Fall alle Materialien zusätzlich in Papier- oder Folienform vorliegen müssten.

Zusammenfassend lassen sich aus den verschiedenen Ausfallszenarien drei Lösungsvorschläge aufzählen:

- Master-CD
- Redundante Server mit dem Lernmanagementsystem am Server- und am Veranstaltungsort
- Alle Materialien in Papier- und Folienform

Die Lösungsvorschläge lassen sich hinsichtlich ihrer Machbarkeit bewerten.

Die Erstellung einer Master-CD bedeutet einen einmaligen Aufwand und ist dann immer verfügbar. Dasselbe gilt auch für den Ausdruck der Materialien.

Weit mehr Aufwand wird durch das Bereitstellen von redundanten Systemen erzeugt. Neben einem fixen Aufwand für die zusätzliche Hardware steht ein beträchtlicher Administrationsaufwand. Dieser Aufwand ist bei der zentralen Bereitstellung mehrerer Server noch zu verkraften. Bei einer dezentralen Lage der LMS-Server kommen die gleichen Kosten wie bei der Aufstellung eines zentralen Server auf den Veranstaltungsort zu. Es sei denn, die dezentralen, in verschiedenen Netzen platzierten Server werden zentral verwaltet. Somit ist am Veranstaltungsort kein zusätzlicher Administrationsaufwand notwendig.

Damit lässt sich als Notfallplan vorschlagen:

Das Lernmanagementsystem sollte mit redundanten, zentralen und dezentralen Servern betrieben werden, und für den Totalausfall der Netze sollte eine Master-CD erstellt werden. Für den Fall ohne lokale Rechner sollten die Materialien in Papier- /Folienform bereitgestellt werden.

Die beiden lokalen Lösungsvorschläge, Master - CD und Ausdruck, sind jedoch nur ein Notbehelf für den einmaligen Einsatz, weil die gesamte Kommunikation über den LMS-Server abläuft. Ohne diesen Server müsste die Veranstaltung anders konzipiert werden.

Auf der Master-CD sollte sich Folgendes befinden:

- sämtliche Materialien zu der Veranstaltung in einer Dateistruktur abgelegt und zum anderen in einer dem Lernmanagementsystem gleichenden verlinkten Form
- das Statistikkabrio einer Version, die direkt von der CD aus ausgeführt werden kann, damit Installationsprobleme auf den Rechnern entfallen.

Gegen den Einsatz einer Master-CD könnte die dynamische Entwicklung von Veranstaltungsinhalten während der Durchführung sprechen.

## J Zielkompetenzen und Instrumente zur Zielerreichung

Der Einsatz multimedialer und anderer Veranstaltungskomponenten verfolgte das Ziel, die Statistikausbildung zu verbessern. Den Studierenden sollten statistische Kompetenzen vermittelt werden, die die bisherige Statistikausbildung nicht oder nur zum Teil vermitteln konnte. Die im folgenden zusammengestellten Fähigkeiten enthalten diese statistischen Kompetenzen, stellen das Ziel unseres Einsatzes neuer Medien in der Statistikausbildung dar. Ob diese erreicht wurden, mit Hilfe welcher Komponenten und mit welchen Formen des Einsatzes, wurde im Rahmen der 2. Durchführung der Evaluationsstudie empirisch überprüft. Die erwünschten Fähigkeiten stellen somit einen

erwünschten Sollzustand der Studierenden als Output der Statistikausbildung dar.

Im Rahmen dieser Arbeit wurden acht zentrale Fähigkeiten herausgearbeitet und nach dem folgenden Schema diskutiert:

1. Es erfolgt eine exakte Beschreibung des gezielten Einsatzes der Komponenten.
2. Zudem wird begründet, warum gerade die jeweils beschriebenen Einsatzvorschläge zu den erwünschten Eigenschaften führen sollten.
3. Außerdem wird diskutiert, mit Hilfe welcher Methoden an welchen Variablen überprüft werden soll, ob die Komponenten tatsächlich zur Zielerreichung führen (=Überprüfung der internen Validität).

### **Verständnis für gezielten und richtigen Einsatz statistischer Methoden zur Problemlösung - Anwendungsverständnis**

⇒ Statistiklabor, statistische Reports, Aufgaben

1. und 2.:<sup>324</sup> Der erste Kontakt der Studierenden mit dem Statistiklabor findet in der Vorlesung statt. Anhand vorbereiteter Laborseiten wird ein reales Problem im Labor analysiert. Statistische Methoden werden gezielt eingesetzt, um ein konkretes Problem zu lösen. Die ersten Vorlesungen und Übungen setzen das Labor aktiv bei der univariaten Datenanalyse ein. Dies geschieht zum einen als reine Präsentation, zum anderen in Form einer gemeinsamen Erstellung von Laborseiten. Für die Übung lösen die Studierenden Aufgaben im Team. Jeder bringt seine Ideen ein, diese werden diskutiert und die statistischen Methoden durchdacht auf die Probleme angewendet. Außerdem soll das Teamwork im Hinblick auf das spätere Berufsleben trainiert werden. Jeder Laboreinsatz findet im Rahmen der Lösung eines realen Problems statt, d.h. ein zielgerichteter Einsatz statistischer Methoden im Labor wird vorgestellt und trainiert. Schrittweise wird der Studierende an die eigene Lösung statistischer Probleme im Statistiklabor herangeführt, zuerst durch die Benutzung des Wizards (Auswählen bereitgestellter statistischer Methoden), dann durch die Benutzung bereitgestellter Bibliotheken (Umgang mit ersten R-Befehlen). In der Übung/Rechnerstunde werden Schwierigkeiten mit der Benutzung des Labors ausgeräumt und diskutiert. Die Übung wird als Diskussionsforum auf Basis von in Übungsaufgaben gestellten Probleme betrachtet. Die Erstellung eines statistischen Reports wird in der Vorlesung vorgestellt. Dabei werden statistische Methoden begründet zur Lösung eines Problems angewendet. Die von den Studierenden erstellten statistischen Reports werden kritisch hinterfragt bezüglich der Anwendung bestimmter Methoden. Dies muss auch innerhalb der Reports geschehen, wie z.B. *das Histogramm erscheint die geeignete grafische Darstellung zu sein, da es sich um ein stetiges Merkmal handelt*. Die Studierenden lernen, die Methoden gezielt auszuwählen und einzusetzen. Die Reports enthalten keine konkreten Aufgabenstellungen, wie *Zeichnen Sie ein Histogramm*. Der Report enthält eine Problemstellung, die es mit Hilfe geeigneter statistischer Methoden zu lösen gilt. Gegebenenfalls enthält die Problemstellung Tipps, die die Studierenden leiten.
3. Während in der Vorlesung trainiert wird, Methoden gezielt einzusetzen, findet eine explorative Evaluation durch teilnehmende Beobachtung<sup>325</sup> statt. Ein Problem wird gestellt, mögliche Methoden diskutiert, dann das Labor zur Problemlösung eingesetzt. Die Aufgaben trainieren dies auch, da die Aufgabenstellungen nach einigen Wochen allgemein gehalten werden, die Studierenden statistische Methoden zur Aufgabenlösung selbst auswählen müssen. Die Beobachtung findet hierbei zum einen über die abgegebenen Aufgaben und zum anderen über die Präsentation der Aufgaben in der Übung und die anschließende kritische Diskussion statt, d.h. der Dozent/Tutor tritt als teilnehmender Beobachter auf. Durch die statistischen Reports wird überprüft, ob die Studierenden in der Lage sind, nur die statistischen Methoden anzuwenden, die das gestellte Problem lösen. Wenden sie alle ihnen bekannten Methoden

<sup>324</sup>Differenzierung durch Natur der Sache nicht möglich.

<sup>325</sup>Wie dabei vorgegangen wird, welche Probleme z.B. bei der Protokollierung auftreten, wird in einem anderen Unterkapitel des Teils IV diskutiert.

an oder genau die, die in der vorherigen Vorlesung vermittelt wurden, wurde das Ziel nicht erreicht. Außerdem ist ihre Begründung für die Auswahl der Methoden entscheidendes Kriterium zur Messung ihres Verständnisses des gezielten Einsatzes statistischer Methoden.

### **Verständnis für Zusammenhänge und Abhängigkeiten zwischen den Methoden**

⇒ Concept Maps, Theoriekomponenten, Glossar, Statistiklabor

1. und 2.<sup>326</sup> Nach Abschluss der Vermittlung einiger statistischer Gebiete (z.B. univariate Datenanalyse, bivariate Datenanalyse) werden in einer Concept Map die Zusammenhänge aller vermittelten statistischen Methoden grafisch dargestellt. Die Concept Map wird in der Vorlesung von dem Dozenten und den Studierenden gemeinsam und aktiv erstellt. Bestehende Zusammenhänge werden diskutiert und eventuelle Fehleinschätzungen der Studierenden beseitigt. Anschließend wird die Theoriekomponente *Zusammenhänge*, die genau diese Zusammenhänge darstellt, vom Dozenten präsentiert. Zusätzlich zeigt der Dozent, dass bestehende Zusammenhänge zwischen den statistischen Methoden durch zahlreiche Verlinkungen der Theoriekomponenten untereinander dargestellt werden. Auch im Glossar werden die bestehenden Zusammenhänge verdeutlicht, da zwischen zusammenhängenden Methoden/Begriffen verlinkt wird. Außerdem kann durch die direkte Gegenüberstellung einer Methode in der Theoriekomponente (z.B. Berechnung des Schätzers ( $b$ )) und eines Glossareintrages (z.B. der Formel zur Berechnung des Korrelationskoeffizienten von Bravais-Pearson) deren Zusammenhang erläutert werden. Im Laufe des Semesters wächst die Concept Map schrittweise. Jedes neu vermittelte Gebiet wird in die bestehende Concept Map aufgenommen. Auf diese Weise erhält der Studierende einen Überblick über die Zusammenhänge aller innerhalb des Semesters vermittelten statistischen Methoden.

Auch das Statistiklabor wird eingesetzt, um Zusammenhänge zwischen statistischen Methoden zu vermitteln. Durch die Anwendung mehrerer statistischer Methoden in kurzen Zeitabständen verstehen die Studierenden die bestehenden Verbindungen. Bei einigen Methoden wird der Zusammenhang auch durch die direkte Verknüpfung der einzelnen Laborobjekte verdeutlicht. (An dieser Stelle muss das Statistiklabor allerdings noch deutlich verbessert werden, um die tatsächlichen Zusammenhänge der Methoden widerzuspiegeln. So wird z. B. ein Stabdiagramm nicht im Labor aus der Häufigkeitstabelle entwickelt, sondern aus der Urliste, was einen falschen Zusammenhang suggeriert.) Durch die Möglichkeit der schnellen Anwendung statistischer Methoden im Labor kann der Fokus der Studierenden von der Erstellung bis zu den Ideen der Methoden und vor allem deren Zusammenhängen untereinander verschoben werden. Außerdem ermöglicht die kompakte grafische Darstellung einen besseren Überblick über die verwendeten Methoden.

3. Der Prozess der in den Vorlesungen gemeinsam erstellten Concept Maps bzw. der der schrittweise erweiterten Concept Map spiegelt wider, ob bestehende Zusammenhänge statistischer Methoden erkannt wurden. Außerdem werden zu bestimmten Themengebieten Aufgaben gestellt, die die Erstellung kleiner Concept Maps erfordern. Die anschließende Diskussion in der Übung, sowie die abgegebenen Lösungsvorschläge spiegeln den Kenntnisstand der Studierenden wider. Des Weiteren wird eine Messung durch qualitative Fragen im Fragebogen wie z.B. *Welcher Zusammenhang besteht Ihrer Meinung nach zwischen den Maßzahlen und den Parametern der Normalverteilung* durchgeführt. Fragen dieser Art können durch quantitative Fragen nicht angemessen beantwortet werden.

### **Verständnis für mathematische Konzepte und schrittweise Herleitung**

⇒ entwickelte Lernmodule zur Varianzzerlegung und zum Gesetz der großen Zahlen, Theoriekomponenten, Skript/Vorlesungsfolien, Statistiklaborseiten, Glossar

1. Zum einen enthalten die Theoriekomponenten mathematische Konzepte. Diese werden in der Vorlesung durch das Skript/Vorlesungsfolien ergänzt und durch Beispiele (Beispielkomponente und Tafel) praktisch erklärt. Für das Gesetz der großen Zahlen und die Varianzzerlegung wurden spezielle Lernmodule angefertigt, die eingesetzt werden, um die formalen

<sup>326</sup>Differenzierung durch Natur der Sache nicht möglich.

Inhalte besser vermitteln zu können. Die Idee besteht darin, komplexe mathematische Beweise zu zerlegen und diese Beweisteile grafisch zu illustrieren. Abschließend erfolgt eine grafische Illustration des Beweises als Einheit. Die Schritte des Beweises befinden sich in einzelnen Theoriekomponenten. Diese sind jeweils verlinkt mit den Statistiklaborseiten, die die zugehörigen grafischen Illustrationen enthalten. Die Zerlegung des Beweises könnte auch ohne Theoriekomponenten stattfinden. Deren Vorteil besteht darin, dass sie untereinander und mit den zugehörigen Statistiklaborseiten verlinkt werden. Es entsteht eine Einheit in Form eines Lernmoduls.

2. Durch den kombinierten Einsatz mehrerer Komponenten (Theorie, Labor, Beispiel, Glossar) zur Vermittlung einer mathematischen statistischen Methode entsteht eine Einheit. Dies wird durch die multimediale Umsetzung erreicht. Hyperlinks verbinden die Komponenten sinnvoll miteinander, man navigiert in nur einem Lernmodul. Die einzelnen Beweisschritte der Theoriekomponenten sind miteinander verlinkt. Der Studierende lernt schrittweise, verliert aber nie den Überblick. Die Concept Map verschafft ihm einen weiteren Überblick über die Einheit *Varianzanalyse* bzw. die Einheit *Gesetz der großen Zahlen*. Bei dem Einsatz in der Statistikvorlesung wird mit der Concept Map begonnen, um einen ersten Überblick über die Vorgehensweise zu geben. Abschließend werden die multimedialen Komponenten in ihrer vorgegebenen Kombination eingesetzt.
3. Die Überprüfung des Erfolgs der entwickelten und eingesetzten Lernmodule erfolgt über Minute Papers, die zum Abschluss der Veranstaltung von den Studierenden einzeln verfasst werden. Deren Thema wird z.B. wie folgt vorgegeben: (Varianzzerlegung) *Betrachten Sie die folgende Situation: Ein Mathetest wird gestellt. Die Ergebnisse streuen in einem großen Intervall. Erläutern Sie, warum die Zusatzinformation MatheLK ja oder nein hilfreich sein kann.*

### **Fähigkeit, statistische Reports zur Dokumentation des statistischen Arbeitens zu verfassen**

⇒ Statistiklabor, Berichterstellungsfunktion, Coachingkomponente, Minireports

1. und 2.<sup>327</sup> Statistische Reports (Idee-Aufbau-Erstellung) stehen im Mittelpunkt der Veranstaltung. In der Vorlesung wird die Coachingkomponente zur Reporterstellung eingeführt. Die in den Lernmodulen enthaltenen statistischen Reports werden in der Vorlesung diskutiert oder zur Nacharbeitung angegeben. Die Studierenden fangen in der dritten Vorlesungswoche selbst an, kleine statistische Reports, sogenannte Minireports, zu verfassen. Diese werden in der Übung kritisch diskutiert.<sup>328</sup> Auf diese Art werden die Studierenden auf die drei großen Reports vorbereitet. Der erste Report wird in der vierten Vorlesungswoche geschrieben. Die Studierenden werden früh an das Reportverfassen herangeführt, um während des Semesters schrittweise Verbesserungen erzielen zu können. Die Komponente *Statistiklabor* wird in jeder Vorlesung und Übung eingesetzt und, wenn sinnvoll, mit der Berichterstellungskomponente kombiniert, um die geeignete Dokumentation statistischen Arbeitens zu präsentieren und zu trainieren. Die beiden ersten Reports werden im Zweierteam angefertigt, damit sich in der Diskussion Lösungsalternativen ergeben, die Diskussion trainiert und anschließend dokumentiert wird. Der dritte Report überprüft in der letzten Vorlesungswoche Einzelleistungen, um genauere Aussagen über die erlangten Fähigkeiten des einzelnen Studierenden bezüglich statistischen Reportverfassens treffen zu können. Das Verfassen statistischer Reports wird also zum einen durch Präsentation und Erklärung, zum anderen durch Aktivität oder "Learning by doing" vermittelt. Kritische Diskussionen der Minireports sowie der großen Reports führen den Studierenden Fehler und Leistungen vor Augen.
3. Die Fähigkeit, statistische Reports zu verfassen, wird zum einen durch die abgegebenen Minireports, zum anderen durch die drei Reports, die die Prüfungsleistung darstellen, gemessen.

<sup>327</sup> Differenzierung durch Natur der Sache nicht möglich.

<sup>328</sup> In einem Unterkapitel des Teils III wird über die Schwierigkeiten des teilnehmendes Beobachtens bezüglich des Reportverfassens berichtet. Es wird ein Konzept für die Reportbewertung erarbeitet.

Auch innerhalb der Diskussionen der Reports in Übung und Vorlesung kann durch die Fragen der Studierenden deren Verständnis eingeschätzt werden.

### **Souveräner Umgang mit neuen Medien und Verständnis für die notwendige Zusammengehörigkeit von "Computational" und "Statistics"**

⇒ *Statistik interaktiv komplett*, Blackboard, Leitfaden

1. Der Umgang mit Neuen Medien wird zum einen durch die Lernsoftware trainiert, zum anderen wird die multimediale Lernsoftware *Statistik interaktiv komplett* im Rahmen des Lernmanagementsystems Blackboard so eingesetzt, dass Aktivität der Studierenden während des Lernprozesses notwendig ist. Zum Beispiel findet das Versenden angefertigter Reports direkt im Lernmanagementsystem statt. Außerdem werden Statistiklaborseiten untereinander über das Lernmanagementsystem ausgetauscht. Die gesamten Veranstaltungsmaterialien stehen nur im Lernmanagementsystem bereit. Die Vorlesungen und Übungen finden grundsätzlich durch Einsatz des Rechners statt. Der Leitfaden sowie das Skript werden im Lernmanagementsystem aufgerufen. Der Leitfaden ist mit multimedialen Komponenten verlinkt, die Vorlesung findet hauptsächlich am Rechner statt. Hier sei angemerkt, dass das Lernmanagementsystem außer der genannten Möglichkeit, die Studierenden durch Aktivität an die neuen Medien heranzuführen, auch Nachteile mit sich bringt.<sup>329</sup> Fazit: Das Lernmanagementsystem und die Lernsoftware allgemein trainieren den Umgang mit den neuen Medien. Das Statistiklabor als eine spezielle Komponente der Lernsoftware verdeutlicht den Studierenden die notwendige Zusammengehörigkeit von "Computational" und "Statistics".
2. Durch die Bereitstellung der Materialien und die Durchführung aller Transaktionen im Lernmanagementsystem müssen sich die Studierenden kontinuierlich mit diesem auseinandersetzen. Durch den Leitfaden werden sie zu relevanten Komponenten der Lernsoftware geführt, lernen diese schrittweise kennen, navigieren selbstständig. Durch die gemeinsame Erstellung von Laborseiten in Vorlesung und Übung, die Vermittlung statistischer Methoden anhand von Laborseiten werden die Studierenden auf die Rolle des Rechners in der Statistik aufmerksam gemacht. Sie lernen direkte Vorteile (Erleichterungen bei der Erstellung grafischer Darstellungen und der Analyse großer Datensätze) und Nachteile (Vermittlungsproblem: keine manuellen Fähigkeiten mehr) kennen. Besonders durch die eigene Aktivität im Labor beim Lösen von Aufgaben, Verfassen von Reports erlernen sie einen souveränen Umgang mit den neuen Medien und erkennen, wieviel Arbeit ihnen das Statistiklabor/der Rechner beim Lösen statistischer Probleme abnehmen kann.
3. Im Rahmen der Veranstaltung können beobachtend erste Schwierigkeiten der Studierenden beim Umgang mit den neuen Medien festgestellt werden - Evaluation durch teilnehmendes Beobachten. Des Weiteren werden die Studierenden in den Fragebögen um Selbsteinschätzungen in diesem Zusammenhang gebeten. Eine weitere Möglichkeit stellt die Auswertung der in Blackboard gegebenen Nutzungsprofile da.

### **Fähigkeit, organisiert und selektiv zu lernen, Prioritäten beim Lernen zu setzen**

⇒ Komponenten der multimedialen Lernsoftware *Statistik interaktiv komplett*, Lernmanagementsystem Blackboard, Leitfaden

1. und 2.<sup>330</sup> Alle Lernmaterialien werden im Lernmanagementsystem organisiert bereitgestellt. Der Leitfaden organisiert die in einer Vorlesung verwendeten Materialien und präsentiert eine Form des organisierten Vermittelns, die der Studierende analog auf sein Lernen übertragen kann.<sup>331</sup> Die große Anzahl verschiedener Medien, die bei der Vermittlung der Statistik eingesetzt werden, fordern den Studierenden auf zu selektieren, welche er in seinem Lernprozess einsetzen möchte, welche ihn beim Lernen am besten unterstützen.

<sup>329</sup>Die Diskussion dieser findet in einem anderen Unterkapitel des Teils IV statt

<sup>330</sup>Differenzierung durch Natur der Sache nicht möglich.

<sup>331</sup>In einem anderen Teil des Teils IV wird die Qualität eines Leitfadens allgemein überprüft - was ist ein guter/schlechter Leitfaden?



Kritisch muss hier die eventuell auftretende Überlastung des Studierenden durch die Vielzahl bereitgestellter Medien betrachtet werden. Der Studierende ist zunächst passiv, indem er das organisierte Bereitstellen von Lernmaterialien im Lernmanagementsystem Blackboard erfährt. Seine Beobachtungen sollen ihn zur Aktivität anregen. Er soll überzeugt werden, seinen eigenen Lernprozess ähnlich zu organisieren. Die Fähigkeit, organisiert und selektiv zu lernen, erfordert diese Aktivität.

3. Die Studierenden werden befragt, welche der Komponenten sie nutzen, wie sie das Lernmanagementsystem bewerten, wie sie mit der Fülle unterschiedlicher Materialien/Komponenten umgehen. Des Weiteren werden sie in den Rechnerstunden und der Übung teilnehmend bezüglich ihres Umgangs mit den einzelnen Komponenten beobachtet.

### **Erster Überblick über Anwendungsfelder, Relevanz und Rolle der Statistik in der Realität**

⇒ Vorlesung, Aufgabensammlung, Reports, Generalbeispiel, Videofallstudie, Verlinkungen zu Statistik-Webpages in BB (Web-Based Statistics Artikel I)

1. Die in der Vorlesung vermittelten Inhalte werden durch die Analyse realer Datensätze im Statistikkolabor illustriert. Es werden Datensätze verwendet, die das Interesse der Studierenden wecken, z.B. Daten, die an ihnen selbst erhoben wurden. Es werden außerdem Datensätze verwendet, die aktuell sind (z.B. EUROSTAT-Datensätze). Sowohl das Generalbeispiel als auch die Videofallstudie basieren auf realen Datensätzen. Die Beispiele und Aufgaben greifen diese auf und sorgen dafür, dass die Studierenden auch selbst reale Problemstellungen behandeln. Die statistischen Reports basieren immer auf realen aktuellen Datensätzen. Des Weiteren werden im Lernmanagementsystem Blackboard Hyperlinks zu Internetseiten statistischer Organisationen gesetzt und am Ende einiger Vorlesungen auf diese verwiesen.
2. Die Vermittlung der statistischen Methoden auf Basis realer Datensätze soll das Interesse der Studierenden für die Statistik wecken und ihnen vor allem vor Augen führen, wie aktuell und wie direkt anwendbar statistische Kenntnisse sind. Sie werden auf statistische Organisationen und deren Aktivitäten aufmerksam gemacht, um die Rolle der Statistik in der Gesellschaft und ihre Relevanz in den Vordergrund zu heben.
3. Zu Beginn und am Ende des Semesters werden die Studierenden gebeten, Statistik und ihre Rolle zu beschreiben. Anhand der Veränderung der Antworten kann gemessen werden, ob das Ziel erreicht wurde, einige Anwendungsfelder der Statistik zu beleuchten und die Rolle der Statistik in der Gesellschaft zu vermitteln.

## **K Diskussion der Arbeitshypothesen des Evaluationsinstrumentes**

In dem ersten Bereich **Typisierung** wurden anstelle von Arbeitshypothesen deskriptive Fragestellungen aufgeworfen:

1. Wie sieht die technische Ausstattung von Teilnehmern einer multimedialen Statistikveranstaltung aus? (Fragebogen1: Fragen 1-3)
  - Die zugehörigen Fragen werden deskriptiv ausgewertet.
  - Die Fragen ermitteln einen Überblick über die Ausstattung der Studierenden. Zusätzlich sind sie Eisbrecherfragen.
2. Wie sieht die Internetnutzung der Teilnehmer einer multimedialen Statistikveranstaltung allgemein und die Nutzung von Blackboard speziell aus?  
Die unterschiedliche Internetnutzung der Teilnehmer wird deskriptiv dargestellt. (Fragebogen1: Fragen 4-6)
3. Wie charakterisieren die Studierenden das Fach Statistik? (Fragebogen1: Fragen 38-42)

- Die zugehörigen Fragen werden deskriptiv ausgewertet.
  - Die Fragen ermitteln einen Überblick über die Einschätzungen der Studierenden.
4. Wie bewerten und charakterisieren die Studierenden die eingesetzte Lernsoftware *Statistik Interaktiv komplett* (Fragebogen1: Fragen 43-47)
- Die zugehörigen Fragen werden deskriptiv ausgewertet.
  - Die Fragen ermitteln einen Überblick über die Einschätzungen der Studierenden in Bezug auf Eigenschaften und Adaptivität der Lernsoftware *Statistik Interaktiv komplett*.
- **Erstellung statistischer Reports, Statistisches Verständnis, Methodenverständnis**
- $H_1$  : Die Vermittlung der Ideen der mathematischen Seite der statistischen Methoden ist durch die neuen Medien besser möglich. (nicht in Fragebögen enthalten)
- Durch Minute Papers wird überprüft, ob die statistischen Methoden, z.B. die Herleitung und Zusammensetzung der Varianzzerlegung, bis zum Ende der Vorlesung verstanden wurden. Die zugehörige Aufgabenstellung lautet: „*Betrachten Sie die folgende Situation: Ein Mathetest wird gestellt. Die Ergebnisse streuen in einem großen Intervall. Erläutern Sie, warum die Zusatzinformation MatheLK ja oder nein hilfreich sein kann.*“
  - Es wird qualitativ evaluiert, da eine subjektive Selbsteinschätzung eigener Fähigkeiten keine Methode sein kann, auf deren Basis argumentiert wird. Die Auswertung der Antworten in den Minute papers gewährleisten eine angemessene objektive Beurteilung.
- $H_2$  : Die Idee sowie das Verfassen statistischer Reports kann durch die neuen Medien besser vermittelt werden.
- Durch Minute Papers wird überprüft, ob die statistischen Methoden, z.B. statistisches Reportverfassen, bis zum Ende der Vorlesung verstanden wurden. Die zugehörige Fragestellung lautet: *Skizzieren Sie die Idee des statistischen Reports.*
  - Es wird qualitativ evaluiert, da eine subjektive Selbsteinschätzung eigener Fähigkeiten keine Methode sein kann, auf deren Basis argumentiert wird. Die Antworten in den Minute papers gewährleisten eine angemessene objektive Beurteilung.
- $H_3$  : Die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik wird durch die Multimediasoftware *Statistik interaktiv komplett* während des Semesters signifikant verändert. (Fragebogen1: 8-9, Fragebogen2: 1-2)
- Es erfolgt eine Einstellungsmessung über das Rosenberg-Modell, d.h. es wird gemessen, ob und wie sich die Einstellung der Studierenden bezüglich der Realitätsnähe und aktuellen Bedeutung der Statistik im Laufe des Semesters verändert hat. Über einen zweiseitigen T-Test wird überprüft, ob eine signifikante Veränderung stattgefunden hat.
  - Das Ziel besteht darin, die Einstellungsveränderung der Studierenden zu überprüfen und festzustellen, welche getroffenen Maßnahmen diese bewirkt haben. Hat z.B. das Reportverfassen zu einer positiven Einstellungsveränderung bei den Studierenden geführt? Wird im Mittel durch den T-Test eine signifikante Veränderung bestätigt, werden in nächsten Schritten mögliche Korrelationen mit anderen Fragen der Fragebögen untersucht. Auf diese Weise kann bestimmt werden, von welchen Maßnahmen die Einstellung wie verändert wurde, d.h. es findet eine Überprüfung der Wirksamkeit der Maßnahmen auf Basis der Einstellung der Studierenden statt.
- $H_4$  : Den Studierenden wird vermittelt, dass es sich bei der Statistik um eine fächerübergreifende Disziplin handelt, die in vielen Fächern eine wichtige Rolle spielt. (Fragebogen1: Fragen 10-11, Fragebogen2: Frage 3)

- Zunächst wird deskriptiv evaluiert, die relativen Häufigkeiten werden grafisch dargestellt. Anschließend werden die Antworten in einer 2x2 Kontingenztabelle erfasst, um einen Chi-Quadrat-Test durchzuführen.
  - Anhand der relativen Häufigkeiten wird ein Überblick darüber gewonnen, ob die Vermittlung erfolgreich war. Mit Hilfe des Chi-Quadrat-Tests wird überprüft, ob der Zeitpunkt der Befragung und die Antwort, ob es sich bei der Statistik um eine fächerübergreifende Disziplin handelt, unabhängig voneinander sind. Besteht keine Unabhängigkeit, so kann daraus geschlossen werden, dass die Veränderung der Bewertung aus Maßnahmen des Einsatzkonzeptes resultiert. Dabei muss die Richtung der Veränderung berücksichtigt werden, dies kann deskriptiv untersucht werden.
- H*<sub>5</sub> : Statistische Reporterstellung kann im Team besser erlernt werden als in Einzelarbeit. (Fragebogen2: Frage 4)
- In einem ersten Schritt wird deskriptiv ein Meinungsbild der Studierenden erstellt. Mit Hilfe der Regressionsanalyse wird der Zusammenhang der Merkmale „Teamarbeit beim statistischen Report verfassen“ und „Vorbereitung auf den ersten/dritten Report“ untersucht.
  - Es soll evaluiert werden, ob das statistische Arbeiten durch Teamarbeit nach Einschätzung der Studierenden gefördert wird.
- H*<sub>6</sub> : Die Fähigkeit, statistische Reports zu verfassen, wird durch den kontinuierlichen Einsatz des Statistiklabors in der Vorlesung und Übung verbessert. (Fragebogen2: Frage 5)
- H*<sub>7</sub> : Die Fähigkeit, statistische Reports zu verfassen, wird durch die Lösung von Aufgaben im Statistiklabor verbessert. (Fragebogen2: Frage 6)
- H*<sub>8</sub> : Die Fähigkeit, statistische Reports zu verfassen, wird durch die Erstellung von Mini-reports innerhalb der ersten Wochen des Semesters verbessert. (Fragebogen2: Frage 7)
- H*<sub>9</sub> : Die Fähigkeit statistische Reports zu verfassen, wird durch die dreimal im Semester stattfindende bewertete Reporterstellung verbessert. (Fragebogen2: Frage 8)
- Um die zeitliche Entwicklung (von Semesterbeginn bis Semesterende) des Umgangs mit dem Statistiklabor bewerten zu können, werden vier Fragen zusammen betrachtet. Diese gehen auf unterschiedliche Zeitpunkte ein und nehmen auf den Umgang mit dem Statistiklabor Bezug. Um zu überprüfen, ob die Einschätzungen der Studierenden zu den unterschiedlichen Zeitpunkten aussagekräftig sind, wird für jeden Studierenden eine Zeitreihe gebildet (auf der x-Achse werden die Fragen abgetragen, auf der y-Achse die Ausprägungen). Anhand der Grafik können die im Semester getroffenen Maßnahmen bewertet werden. War es z.B. für die Erstellung des Abschlussreports hilfreich, zuvor Mini-reports anzufertigen?  
Die Zusammenhänge der unterschiedlichen Maßnahmen des Einsatzkonzeptes werden per Regressionsanalyse und Berechnung des Korrelationskoeffizienten von Bravais-Pearson untersucht.
  - Es sollen Verbesserungsmöglichkeiten des Einsatzkonzeptes durch die Evaluation gesucht werden. Es kann z.B. sein, dass das im Einsatzkonzept vorgesehene "Training" (Bearbeitung von Aufgaben, Erstellung von Mini-reports) keinen wesentlichen Einfluss auf die Fähigkeiten zur Erstellung des Abschlussreports haben, d.h. die dem Einsatzkonzept zugrunde liegende Einschätzung nicht der Realität entspricht. Zu beachten ist hierbei, dass es sich um Einschätzungen der Studierenden handelt. Diese könnten durch die Maßnahmen einen Lernerfolg erzielt, dies jedoch nicht wahrgenommen haben (subjektive Fehleinschätzung).
- **Adaptivität (Nutzungsgrad-Nutzungsdauer, Akzeptanz, Interaktivität) der multimedialen Komponenten**

$H_{10}$  : Es besteht ein Zusammenhang zwischen verschiedenen Komponenten des Einsatzkonzeptes der multimedialen Statistikveranstaltung bezüglich der Eigenaktivität der Studierenden (Fragebogen1: Fragen 12-16, Fragebogen2: Frage 9)

- Die Eigenaktivität der Studierenden wird deskriptiv bezüglich der Aspekte Lösung zusätzlicher Übungsaufgaben, Betrachtung zusätzlicher multimedialer Komponenten, Aktivität in der Vorlesung und eigenständiges Arbeiten im Statistiklabor analysiert. In einer mehrdimensionalen Kontingenztafel werden mögliche Zusammenhänge zwischen diesen Merkmalen untersucht, z.B. der Zusammenhang zwischen der Aktivität in der Vorlesung und der Lösung zusätzlicher Übungsaufgaben. Außerdem wird untersucht, ob die Studierenden Eigenaktivität zeigen, indem sie das Statistiklabor auf Bereiche außerhalb der Veranstaltung anwenden. Hier bietet sich eine deskriptive Darstellung der einzelnen Bereiche an.
- Zum einen wird die Qualität und Adaptivität des Statistiklabors evaluiert. Weiterhin kann die mögliche Anwendung des Statistiklabors in anderen Bereichen zeigen, dass die Studierenden die Interdisziplinarität der Statistik erkannt haben.

$H_{11}$  : Der gezielte Einsatz des Laborobjektes *Kalkulator* ermöglicht eine schrittweise Heranführung an Programmiersprachen zur Verringerung von Barrieren (Fragebogen1: Fragen 17-22, Fragebogen2: Fragen 10-11)

- Zunächst wird deskriptiv untersucht, ob ein erster Kontakt mit Programmiersprachen bereits stattgefunden hat. Mit Hilfe eines Plots wird untersucht, ob ein Zusammenhang zwischen dem Kenntnisstand bezüglich Programmierens und der eingeschätzten Fähigkeit besteht, ein Problem im Statistiklabor zu lösen. Des Weiteren wird deskriptiv untersucht, welche Laborobjekte von den Studierenden präferiert werden, R-Grafik Wizard versus R-Kalkulator. Durch die Anwendung eines zweiseitigen T-Tests wird überprüft, ob im Laufe des Semesters eine Veränderung bezüglich der Einschätzung des R-Kalkulators stattfindet. Ein weiterer zweiseitiger T-Test überprüft die Veränderung beim Erlernen der Anwendung des R-Kalkulators. Die deskriptive Analyse der Nutzung der Onlinehilfe bezüglich verschiedener Laborobjekte zeigt, wo die Präferenzen der Studierenden liegen, im Klicken oder im Programmieren. Außerdem wird der Stellenwert der Onlinehilfe deutlich.
- Die deskriptive Analyse zeigt, wo die Studierenden stehen, d.h. wo mit der Einführung ins Statistiklabor angesetzt werden muss. Das Einsatzkonzept wird wie folgt evaluiert. Durch die Überprüfung des Zusammenhangs mit Hilfe des Interaktionsplots können Aussagen darüber getroffen werden, ob Studierende ohne Programmierkenntnisse genauso gut an das Statistiklabor herangeführt wurden wie Studierende, die erste Kontakte mit dem Programmieren bereits vor der Veranstaltung hatten. Der Vergleich zwischen R-Kalkulator und R-GrafikWizard evaluiert die Adaptivität dieser Laborobjekte. Außerdem kann festgestellt werden, ob die Studierenden tatsächlich an das Programmieren herangeführt werden oder dieses wenn möglich durch den R-GrafikWizard umgehen. Dies ist gleichzeitig eine Evaluation des Einsatzkonzeptes, das das Ziel verfolgt, den Studierenden die Vorteile des Programmierens zu vermitteln. Weicht ein Großteil der Studierenden dem aus, ist der Einsatz falsch konzipiert.<sup>332</sup> Auch die Untersuchung der Einstellungsveränderung zeigt, ob die im Einsatzkonzept vorgesehenen Maßnahmen zur Erhöhung der Adaptivität des R-Kalkulators während des Semesters ihre Ziele erreicht haben. Die Evaluation der Onlinehilfe gibt einen Überblick über ihre Nutzung. Wird diese viel genutzt, kann sie in der Zukunft weiterentwickelt werden.

---

<sup>332</sup>Das Einsatzkonzept sieht vor, dass den Studierenden Hilfen bei der Anwendung des R-Kalkulators (z.B. Beispiellaborseiten) gegeben werden. Dies muss bei der Evaluation berücksichtigt werden. Dies ist zurückzuführen auf den zur Zeit noch mangelnden Reifegrad des Statistiklabors.

- **Einsatzkonzept für multimediale Veranstaltung**

- $H_{12}$  : Die Studierenden erwarten von einer multimedialen Statistikveranstaltung, dass die eingesetzten Komponenten (Lernmanagementsystem, Lernsoftware, Statistiklabor) für sie in naher Zukunft (Studium und Beruf) Relevanz haben. (Fragebogen1: Fragen 23-25)
- Zunächst wird deskriptiv untersucht, ob die Studierenden die Lernsoftware als notwendiges Instrument zur Vermittlung der Vorlesungsinhalte ansehen. Ob die Studierenden in einer multimedialen Veranstaltung eine modernere Vermittlung der Statistik erwarten, wird deskriptiv anhand zweier Aspekte untersucht: erstens der Einschätzung des Einflusses des Statistiklabors auf später notwendige statistisch-methodische Fähigkeiten, zweitens die Einschätzung der Relevanz des Lernmanagementsystems Blackboard im Einsatzkonzept. Anschließend wird mit Hilfe der Regression der Zusammenhang dieser beiden Merkmale analysiert und durch die Berechnung des Korrelationskoeffizienten von Bravais-Pearson numerisch belegt.
  - Die deskriptive Analyse überprüft zunächst die Adaptivität der Lernsoftware generell. Anschließend wird speziell die Adaptivität des Lernmanagementsystems Blackboard und des Statistiklabors evaluiert. Dies bedeutet gleichzeitig eine Evaluation des Einsatzkonzeptes. Die mit dem Lernmanagementsystem erwarteten Vorteile werden überprüft, ebenso die Rolle, die dem Statistiklabor im Einsatzkonzept zukommt. Der Zusammenhang zwischen dem Lernmanagementsystem Blackboard und dem Statistiklabor bezüglich einer moderneren Vermittlung der Statistik wird durch Anwendung der Regression überprüft.
- $H_{13}$  : Die Studierenden erwarten, dass ihnen das Erlernen statistischer Inhalte durch die multimediale Lernsoftware erleichtert wird. (Fragebogen1: Fragen 26-27, Fragebogen2: Fragen 12-13)
- Unter Anwendung des Rosenberg-Modells findet eine Einstellungsmessung zu Beginn und zum Ende des Semesters statt. Durch den Vergleich beider Messungen kann eine mögliche Einstellungsveränderung im Laufe des Semesters festgestellt und mit dem T-Test überprüft werden.
  - Durch die Einstellungsmessung soll überprüft werden, ob die Studierenden durch die multimediale Lernsoftware zum Statistikkennen motiviert werden, da ihnen dies leichter erscheint.
- $H_{14}$  : Das gemeinsame Erstellen von Statistiklaborseiten in der Vorlesung und Übung durch Dozenten und Studierende zeigt deren Möglichkeiten und Arbeitserleichterungen auf und animiert zur Nutzung. (Fragebogen1: Fragen 29-32, Fragebogen2: Frage 14)
- Die Studierenden sollen zu Semesterbeginn und Semesterende ihre Fähigkeiten einschätzen, im Statistiklabor die Vorlesungen nachzuarbeiten. Mit Hilfe eines zweiseitigen T-Tests wird eine Veränderung ihrer Einschätzung untersucht. Des Weiteren wird deskriptiv untersucht, ob die Studierenden selbstständig Statistiklaborseiten erstellen, sie in dem Statistiklabor eine Arbeitserleichterung sehen und ob sie den Eindruck haben, dass das Statistiklabor umfassende statistische Lösungsmöglichkeiten bietet. Die Zusammenhänge der Merkmale werden in einer mehrdimensionalen Kontingenztafel dargestellt und untersucht.
  - Zum einen findet eine Evaluation des Einsatzkonzeptes statt: Bei der Messung einer positiven Veränderung können die im Einsatzkonzept verankerten Maßnahmen zur Erstellung von Laborseiten in der Vorlesung / Übung / Zuhause als erfolgreich bewertet werden. Zum anderen werden einzelne Aspekte der Qualität der Komponente Statistiklabor untersucht. Die Adaptivität, die Arbeitserleichterung und die Komplexität bezüglich der Anwendbarkeit statistischer Methoden.

- **Lernmanagementsystem Blackboard (BB)**

- $H_{15}$  : Die Organisation des Lernens und der Umgang mit neuen Medien wird durch deren Einsatz in einem Lernmanagementsystem verbessert. (Fragebogen1: Frage 34)
- Deskriptiv wird untersucht, ob das Lernmanagementsystem Blackboard den Studierenden eine bessere Organisation ihres Lernens ermöglicht.
  - Es soll herausgefunden werden, ob das Lernmanagementsystem einen wirklichen Vorteil bezüglich der Organisation des Lernens mit neuen Medien darstellt.
- $H_{16}$  : Traditionelle und multimediale Komponenten lassen sich in einem Lernmanagementsystem sinnvoll miteinander kombinieren. (Fragebogen1: Frage 35, Fragebogen2: 16)
- Die Studierenden werden sowohl zu Semesterbeginn als auch zu Semesterende bezüglich der Eignung Blackboards zur Kombination traditioneller und multimedialer Veranstaltungskomponenten befragt. Eine mögliche Veränderung im Mittel wird anhand eines zweiseitigen T-Tests überprüft.
  - Es soll untersucht werden, ob das Lernmanagementsystem tatsächlich geeignet ist, verschiedenste Veranstaltungskomponenten sinnvoll miteinander zu kombinieren. Außerdem wird das Einsatzkonzept evaluiert, d.h. es wird untersucht, ob der konzipierte Einsatz und die Evaluation der multimedialen Lernsoftware im Rahmen eines Lernmanagementsystems die angestrebten Ziele erreicht.
- $H_{17}$  : Die Teamarbeit der Studierenden wird durch das Lernmanagementsystem positiv unterstützt. (Fragebogen1: Fragen 36-37, Fragebogen2: Frage 17)
- Zunächst wird deskriptiv untersucht, wie die Studierenden das Lernmanagementsystem hinsichtlich der Kommunikation untereinander und mit den Dozenten bewerten. Außerdem wird überprüft, wie das Lernmanagementsystem im Vergleich zu alternativen Kommunikationsformen (Webseite/Email) bewertet wird. Eine Veränderung der Bewertung von Blackboard als Kommunikationsplattform wird durch einen T-Test überprüft, auf Basis eines Vergleichs der Ergebnisse zu Semesterbeginn und zu Semesterende. Ein möglicher Zusammenhang zwischen der Bewertung der beiden genannten Aspekte wird im Rahmen einer Regressionsanalyse näher untersucht. Ein mögliches Ergebnis könnte sein, dass sich folgende Tendenz herausbildet: Wird das Lernmanagementsystem als Kommunikationsplattform positiv eingestuft, so werden mögliche Alternativen negativ bewertet.
  - Es soll herausgefunden werden, ob sich der finanzielle Aufwand, der mit dem Einsatz des Lernmanagementsystems Blackboard verbunden ist, gegenüber kostengünstigen Alternativen durch Leistung auszeichnet. Ein entscheidender Aspekt in diesem Zusammenhang ist die Teamarbeit, der im Einsatzkonzept eine zentrale Rolle zukommt. Insofern wird evaluiert, ob die konzipierten Maßnahmen zur Förderung bzw. zum Erlernen und Training der Teamarbeit im Lernmanagementsystem ihr Ziel erreichen, oder ob an dieser Stelle eine Modifikation des Einsatzkonzeptes notwendig wird.