

Evaluation von Augmented Chemistry, einem computergestützten Lernsystem für den Chemieunterricht: eine Vergleichsstudie

Master Thesis

Author(s):

Bötschi, Kristina

Publication date:

2005

Permanent link:

<https://doi.org/10.3929/ethz-a-005053243>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich



Institut für Hygiene und Arbeitsphysiologie, BEPR, ETH Zürich

in cooperation with

Psychologisches Institut der Universität Zürich
Abteilung Angewandte Psychologie
Fachrichtung Psychologische Methodenlehre

Master of Science thesis of Kristina Bötschi

**Evaluation von *Augmented Chemistry*, einem
computergestützten Lernsystem für den
Chemieunterricht: Eine Vergleichsstudie**

**Evaluation of *Augmented Chemistry*, a Tangible
User Interface for high-school organic chemistry
education**

Advisors:

Amos Cohen, University of Zurich
Morten Fjeld, ETH Zurich

Zürich, February 2005

Summary in English:

The comparative study presented here is an evaluation of *Augmented Chemistry (AC)*, an interactive, computer-supported, learning system for organic chemistry. The system is a tangible user interface that aims to enhance and reinforce lessons on organic molecule composition. The study's purpose was to determine whether AC can be successfully used in chemistry lessons at a secondary school level. This empirical evaluation compares learning effectiveness and user acceptance (the dependent variables) of AC versus the *Ball-and-stick Model (BSM)* (the independent variable). It was carried out at a secondary school with twenty-six biology and laboratory students, all in the first year of their secondary education (15-17 years of age). The effect of the learning method (AC or BSM) on the dependent variables was measured repeatedly in an experiment of the AB-BA design. The students were divided into two groups of equal size and worked individually with AC and BSM at two different times. One group used AC in a first session and then the BSM in a second session. The other group used the systems in reverse order. The users' mood, mental task load, physical task load, satisfaction, perceived usability, and system preferences were measured using questionnaires and then converted into an assessment of user acceptance. The results showed that subjects using the BSM solved the problems more effectively than those using AC. However, retention of the lesson, as measured in a subsequent test did not reveal any significant difference. There were also only slight differences in subjective usability between the two methods. System preference was measured by nine subjective criteria. In terms of ease of learning the system, ease of use, helpfulness in problem-solving, and comfort of use, the BSM outperformed AC. In terms of enjoyability, visualization, content availability, future use, and effectiveness of learning, AC outperformed the BSM. In conclusion, this study shows that the tested version of AC does not provide a learning environment superior to that of the BSM. In order to replace the BSM with AC, a higher user acceptance of AC is required. If its usability is improved and its functionality enhanced, AC may potentially provide a superior learning experience to that of the BSM.

Psychologisches Institut der Universität Zürich
Abteilung Angewandte Psychologie
Fachrichtung Psychologische Methodenlehre

Lizentiatsarbeit

**Evaluation von *Augmented Chemistry*, einem
computergestützten Lernsystem für den
Chemieunterricht: Eine Vergleichsstudie**

Referent: Prof. Dr. Amos Cohen
Co-Referent: Prof. Dr. Morten Fjeld

Kristina Bötschi
Salstrasse 7
8400 Winterthur
052 203 36 86
kristina@gehring.biz

Zürich, Februar 2005

Abstract

Mit der vorliegenden Lizentiatsarbeit wurde das interaktive, computergestützte Lernsystem *Augmented Chemistry* (AC) evaluiert. Das auf einer haptischen Benutzerschnittstelle basierende Lernsystem dient dem Lernen im Bereich Molekülbau. Die Evaluationsstudie gibt Aufschluss darüber, ob AC erfolgreich im Chemieunterricht eingesetzt werden kann und beinhaltet einen experimentellen Vergleich von AC mit dem traditionellen Lernmittel *Kugel-Stab-Modell* (KSM), bezüglich Lernwirksamkeit und Benutzerakzeptanz. Die Untersuchung fand in einem Ausbildungszentrum für naturwissenschaftliche Berufe statt und wurde mit 26 Biologielaboranten im ersten Ausbildungsjahr durchgeführt. Die Effekte der Lernmethode auf die abhängigen Variablen wurden durch wiederholte Messungen in einem AB/BA-Design untersucht. Dabei wurden AC und KSM zu zwei Zeitpunkten anhand von zwei Gruppen getestet, welche in unterschiedlicher Reihenfolge mit den beiden Lernmethoden Aufgaben bearbeiten mussten. Mit Fragebögen wurden emotionale Stimmung, Beanspruchung und Zufriedenheit der Lernenden sowie subjektive Usability von AC erfasst, um aus diesen Bewertungen Rückschlüsse auf die Benutzerakzeptanz zu ziehen. Es zeigte sich, dass die Lernenden mit dem KSM Aufgaben signifikant besser lösen konnten als mit AC; im Lernerfolg, welcher über Prüfungen erhoben wurde, waren die Unterschiede jedoch unbedeutend. In den subjektiven Bewertungen der beiden Lernmethoden gab es ebenfalls nur kleine Unterschiede, und die Usability von AC wurde von den Lernenden mittelmässig beurteilt. Infolgedessen konnte sich AC nicht als erfolgreichere Lernumgebung durchsetzen. Um ein traditionelles Lernmittel wie das KSM zu ersetzen sind höhere Lernerfolge und bessere Benutzerakzeptanz notwendig. Durch Optimierung der Bedienbarkeit und Erweiterung der Funktionalität könnte AC das Lernen in Zukunft positiv beeinflussen.

Inhaltsverzeichnis

Abbildungsverzeichnis	V
Tabellenverzeichnis.....	VII
1. Einleitung	1
2. Theoretischer Hintergrund	4
2.1 Lernen mit dem Computer	4
2.1.1 Begriffsbestimmung.....	4
2.1.2 Computergestützte Lernsysteme und ihre Verwendung im Bildungsbereich.....	5
2.1.3 Virtuelle Realität	8
2.2 Psychologische Lerntheorien.....	11
2.2.1 Behaviorismus.....	12
2.2.2 Kognitivismus	12
2.2.3 Konstruktivismus	13
2.3 Wirksamkeit computergestützter Lernsysteme.....	14
2.3.1 Evaluationsforschung im Bildungsbereich.....	15
2.3.2 Messung des Lernerfolgs von computergestützten Lernsystemen	17
2.3.3 Usability	18
2.3.4 Messung der Beanspruchung	20
2.3.5 Empirische Befunde zur Wirksamkeit computergestützter Lernsysteme ...	21
2.3.6 Methodische und praktische Einschränkungen von Evaluationen	23
2.4 Chemie und Computer.....	24
2.5 Das Lernsystem Augmented Chemistry	26
2.5.1 Entstehungsgeschichte	27
2.5.2 Systembeschreibung	27
2.5.3 Funktionsweise und Benutzung	30
2.5.4 Didaktische Motive	31

2.6	Das Lernsystem Kugel-Stab-Modell	33
2.7	Unterschiede und Gemeinsamkeiten von Augmented Chemistry und dem Kugel-Stab-Modell	34
2.8	Abschliessende Bemerkungen	35
3.	Fragestellungen	36
4.	Methode	37
4.1	Einschränkungen für das Augmented Chemistry-System	37
4.2	Stichprobe.....	38
4.3	Versuchsplan	39
4.3.1	AB/BA-Design	41
4.4	Untersuchungsvariablen und ihre Operationalisierung.....	42
4.4.1	Lernwirksamkeit	42
4.4.2	Beanspruchung	45
4.4.3	Emotionale Stimmung	46
4.4.4	Zufriedenheit mit der Lernsituation.....	47
4.4.5	Präferenzbeurteilung	48
4.4.6	Subjektive Usability von Augmented Chemistry	48
4.5	Versuchsdurchführung.....	49
4.6	Auswertungsverfahren.....	53
4.6.1	Auswertung des AB/BA-Designs	53
4.6.2	Zusammenhänge	57
4.6.3	Subjektive Usability	57
5.	Ergebnisse	58
5.1	Allgemeine Zusammenhänge	58

5.2	Lernwirksamkeit.....	60
5.2.1	Verteilung der Leistung in den Lernaufgaben und im Kurztest.....	60
5.2.2	Unterschiede in der Lernwirksamkeit zwischen den beiden Lernmethoden Augmented Chemistry und Kugel-Stab-Modell	67
5.3	Beanspruchung.....	72
5.3.1	Verteilung der Untersuchungspersonen hinsichtlich ihrer Gesamtbeanspruchung	72
5.3.2	Unterschiede in der Beanspruchung zwischen Personen, die mit Augmented Chemistry, und solchen, die mit dem Kugel-Stab-Modell gelernt haben	73
5.4	Emotionale Stimmung.....	77
5.4.1	Verteilung der Untersuchungspersonen hinsichtlich ihrer emotionalen Stimmung.....	77
5.4.2	Unterschiede in der emotionalen Stimmung bei Personen, die mit Augmented Chemistry, und bei Personen, die mit dem Kugel-Stab-Modell gelernt haben	80
5.5	Zufriedenheit beim Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell	84
5.5.1	Unterschiede in der Zufriedenheit beim Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell.....	85
5.5.2	Berichtete positive und negative Aspekte zum Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell.....	86
5.5.3	Benotung von Augmented Chemistry und dem Kugel-Stab-Modell	87
5.6	Präferenz für ein Lernsystem.....	89
5.7	Subjektive Usability von Augmented Chemistry	90
5.8	Zusammenhänge zwischen Lernerfolg und subjektiven Variablen.....	91
5.9	Prädiktive Bedeutung von Lernmethode, Periode, emotionaler Stimmung und Beanspruchung für den Lernerfolg	94
6.	Diskussion.....	95
6.1	Interpretation der Ergebnisse	95

6.1.1	Lernwirksamkeit	95
6.1.2	Beanspruchung	98
6.1.3	Emotionale Reaktion	98
6.1.4	Präferenzbeurteilung	100
6.1.5	Zufriedenheit	101
6.1.6	Akzeptanz: Eine Integration aller untersuchten Aspekte	102
6.2	Zusammenhänge	103
6.3	Subjektive Usability von Augmented Chemistry	104
6.4	Potenzial von Augmented Chemistry als erfolgreiche Lernumgebung	105
6.4.1	Curriculare Einbindung von Augmented Chemistry	106
6.4.2	Didaktische Aspekte von Augmented Chemistry	107
6.4.3	Interaktivitätsgrad von Augmented Chemistry	107
6.5	Probleme und kritische Betrachtungen	108
6.6	Ausblick	109
7.	Zusammenfassung	112
8.	Literaturverzeichnis	115
Anhang	123
Anhang A:	Aufgabenblätter der beiden Versuche	123
Anhang B:	Kurztests der beiden Versuchen	125
Anhang C:	Nasa Task Load Index	126
Anhang D:	Self-Assessment-Manikin (SAM)	129
Anhang E:	Fragebogen zur Zufriedenheit	130
Anhang F:	Fragebogen zur Präferenz eines Lernsystems	132
Anhang G:	Software Usability Measurement Inventory (SUMI)	133

Abbildungsverzeichnis

Abb. 1.1:	Das Lernsystem Augmented Chemistry.....	2
Abb. 1.2:	Das Lernsystem Kugel-Stab-Modell.	2
Abb. 2.1:	Beispiel eines Head-up-Displays im Flugzeugcockpit.	10
Abb. 2.2:	Beispiel eines Head-Mounted-Displays zur Darstellung einer Roboter-Anwendung.	10
Abb. 2.3:	Die Arbeit mit Construct3D, einem kollaborativen AR-System.....	11
Abb. 2.4:	CIELT-Rahmenmodell für die Evaluation von Lernmedien.	16
Abb. 2.5:	Anwendungsrahmen der Gebrauchstauglichkeit nach ISO-Norm 9241-11.	19
Abb. 2.6:	Systemskizze von AC.	30
Abb. 2.7:	AC: Funktionsschritt 1.....	30
Abb. 2.8:	AC: Funktionsschritt 2.....	30
Abb. 2.9:	AC: Funktionsschritt 3.....	30
Abb. 2.10:	Symbolisierung der Bindungsarten bei AC durch das Tetraedermodell.....	31
Abb. 2.11:	Der Molekülbaukasten: Kugel-Stab-Modell (KSM).	34
Abb. 2.12:	Molekülmodelle nach dem Prinzip des KSM.....	34
Abb. 2.13:	Moleküldarstellung: <i>Ball-Stick</i>	34
Abb. 2.14:	Moleküldarstellung: <i>Wireframe</i>	34
Abb. 4.1:	Augmented Chemistry - Arbeitsplatz.....	38
Abb. 4.2:	Das Versuchsdesign.	40
Abb. 4.3:	Musterverlauf einer Lernaufgabe in den Aufgabenblättern.	43
Abb. 4.4:	Bildvorlage im Self Assessment Manikin (SAM) zur Beurteilung der emotionalen Dimensionen <i>Aktiviertheit</i> , <i>Valenz</i> und <i>Dominanz</i>	46
Abb. 4.5:	Räumliche Anordnung der AC-Stationen im Versuchsraum.	51
Abb. 4.6:	Räumliche Anordnung der KSM-Baukästen im Versuchsraum.....	51
Abb. 4.7:	Vergleichende Darstellung des Vorgehens bei der Aufgabenlösung mit AC und mit dem KSM.	52
Abb. 4.8:	Reiner Behandlungseffekt.....	55
Abb. 4.9:	Periodeneffekt mit Behandlungseffekt.	55
Abb. 4.10:	Carry-Over-Effekt.....	56
Abb. 5.1:	Häufigkeitsverteilung der Leistung in den Lernaufgaben über beide Perioden hinweg.	61
Abb. 5.2:	Häufigkeitsverteilung der Leistung im Kurztest über beide Perioden hinweg.	62
Abb. 5.3:	Verteilung der Leistung in den Lernaufgaben getrennt nach AC und dem KSM.	62

Abb. 5.4:	Verteilung der Leistung im Kurztest getrennt nach AC und dem KSM.....	63
Abb. 5.5:	Verteilung der erreichten Punktzahl in den Lernaufgaben getrennt nach AC und dem KSM und separat für beide Perioden (mit Ausreisser!)	64
Abb. 5.6:	Verteilung der erreichten Punktzahl in den Lernaufgaben getrennt nach AC und dem KSM und separat für beide Perioden (ohne Ausreisser!)	64
Abb. 5.7:	Verteilung der erreichten Punktzahl im Kurztest getrennt nach AC und dem KSM und separat für beide Perioden.....	66
Abb. 5.8:	Transformierte Verteilung der Leistung in den Lernaufgaben über beide Perioden hinweg.	68
Abb. 5.9:	Die durchschnittliche Leistung in den Lernaufgaben pro Lernmethode und nach der Periode.....	69
Abb. 5.10:	Die durchschnittliche Leistung im Kurztest pro Lernmethode und nach der Periode.	71
Abb. 5.11:	Häufigkeitsverteilung der Gesamtbeanspruchung bei den Upn über beide Perioden hinweg.	72
Abb. 5.12:	Verteilung der Gesamtbeanspruchung getrennt nach AC und dem KSM.....	72
Abb. 5.13:	Die mittlere Gesamtbeanspruchung pro Lernmethode und nach der Periode.	74
Abb. 5.14:	Die durchschnittlichen Beanspruchungswerte in den sechs Subskalen des NASA-TLX getrennt nach AC und dem KSM.	75
Abb. 5.15:	Häufigkeitsverteilung der Aktiviertheit bei den Upn über beide Perioden hinweg.	78
Abb. 5.16:	Verteilung der Aktiviertheit getrennt nach AC und dem KSM.....	78
Abb. 5.17:	Häufigkeitsverteilung der Valenz bei den Upn über beide Perioden hinweg. ..	79
Abb. 5.18:	Verteilung der Valenz getrennt nach AC und dem KSM	79
Abb. 5.19:	Häufigkeitsverteilung der Dominanz bei den Upn über beide Perioden hinweg.	80
Abb. 5.20:	Verteilung der Dominanz getrennt nach AC und KSM	80
Abb. 5.21:	Die mittlere Aktiviertheit pro Lernmethode und nach der Periode.....	82
Abb. 5.22:	Die mittlere Valenz pro Lernmethode und nach der Periode.	83
Abb. 5.23:	Das mittlere Dominanzempfinden pro Lernmethode und nach der Periode....	84
Abb. 5.24:	Häufigkeitsverteilung der Benotung von AC über beide Perioden hinweg.....	87
Abb. 5.25:	Häufigkeitsverteilung der Benotung vom KSM über beide Perioden hinweg...	87
Abb. 5.26:	Die durchschnittliche Bewertung der Lernmethoden auf einer Notenskala pro Lernmethode und nach der Periode.	88
Abb. 5.27:	Relative Häufigkeiten der Präferenzangaben zu den Aspekten aus dem Fragebogen zur Präferenz eines Lernsystems.	89
Abb. 5.28:	Verteilungen der SUMI-Skalenwerte.....	90
Abb. 6.1:	Befindlichkeiten aus der Kombination der Extremausprägungen in der Valenz- und Aktiviertheitsdimension.	100

Tabellenverzeichnis

Tab. 2.1:	Systemtypen, die auf Virtueller Realität basieren (nach Biocca und Delaney (1995)).	8
Tab. 2.2:	Prinzipien der klassischen Strukturlehre, die mit AC erarbeitet werden können.	32
Tab. 4.1:	Behandlungsabfolge in einem AB/BA-Design.	41
Tab. 4.2:	Die fünf Subskalen von SUMI und ihre Bedeutung.	49
Tab. 4.3:	Faktoren der Varianzanalyse des AB/BA-Designs.	53
Tab. 4.4:	Die Haupteffekte des varianzanalytischen Designs mit ihren entsprechenden Interaktionseffekten.	54
Tab. 4.5:	Interpretationsgrundlage der drei Effekte (Behandlungs-, Perioden-, Carry-Over-Effekt) eines AB/BA-Designs aufgrund der Faktoren der Varianzanalyse.	54
Tab. 5.1:	Korrelationsmatrix der wichtigsten Variablen der Untersuchung.	58
Tab. 5.2:	Korrelationsmatrix nur von der ersten Periode (Thema A).	59
Tab. 5.3:	Korrelationsmatrix nur von der zweiten Periode (Thema B).	59
Tab. 5.4:	Übersicht über die wichtigsten Kennwerte aus Abbildung 5.6.	65
Tab. 5.5:	Übersicht über die wichtigsten Kennwerte aus Abbildung 5.7.	66
Tab. 5.6:	Mittlere Leistung in den Lernaufgaben mit AC und dem KSM getrennt nach der Periode.	68
Tab. 5.7:	Mittlere Leistung in den Kurztests mit AC und dem KSM getrennt nach der Periode.	70
Tab. 5.8:	Mittlere Beanspruchung (gewichteter Gesamtwert) beim Lernen mit AC und mit dem KSM getrennt nach der Periode.	74
Tab. 5.9:	Beanspruchung der Lernenden in den sechs Subdimensionen mit ANOVA-Analyse auf Unterschiede zwischen den Lernmethoden.	77
Tab. 5.10:	Die Mittelwerte in den drei Dimensionen der emotionalen Stimmung für Personen, die mit AC und jene, die mit dem KSM gearbeitet haben getrennt nach der Periode.	81
Tab. 5.11:	Zufriedenheit der Upn beim Lernen mit AC bzw. dem KSM mit ANOVA-Analyse auf Unterschiede zwischen den Lernmethoden.	85
Tab. 5.12:	Lernsituation AC: Korrelationen der objektiven Variablen zur Lernwirksamkeit mit den subjektiven Variablen emotionale Stimmung und Beanspruchung.	92
Tab. 5.13:	Lernsituation KSM: Korrelationen der objektiven Variablen zur Lernwirksamkeit mit den subjektiven Variablen emotionale Stimmung und Beanspruchung.	92
Tab. 5.14:	Ergebnisse der multiplen Regression des Lernerfolgs auf die Variablen Lernmethode, Periode, emotionale Stimmung und Beanspruchung.	94

1. Einleitung

Medien sind im Unterricht und zum Lernen nicht mehr wegzudenken. Technische Medien wie der Computer haben in den letzten Jahren zunehmend an Bedeutung gewonnen. Die Entwicklung neuer computergestützter Lernmedien soll das Lernen erleichtern und unterstützen. Die rasante Entwicklung von Informations- und Kommunikationstechnologien bieten immer neue Möglichkeiten in diesem Bereich. *Multimedia* ist in den letzten Jahren zum neuen Schlagwort in der mediendidaktischen Diskussion geworden und Lernen in *Virtuellen Realitäten* eröffnet neue Perspektiven. Es werden vermehrt sogenannte virtuelle Erlebniswelten geschaffen, welche die Phantasie anregen, Wissen transportieren und Spass am Lernen vermitteln sollen. Der Einsatz Virtueller Realitäten ist gerade bei der Darstellung abstrakter, in der Realität nicht erfahrbarer Informationen und beim Aufzeigen komplexer Zusammenhänge bedeutsam. Eine Reihe neuer Möglichkeiten der Wissensvermittlung hat sich entwickelt. Die Hoffnungen und Erwartungen in solche neuen Technologien im Bildungsbereich sind entsprechend hoch: das Lernen soll schneller, einfacher, effizienter werden. Lernmedien der Zukunft versprechen Bildungsinnovationen oder gar –revolutionen, wobei aber zu einer gewissen Skepsis gemahnt wird (vgl. Kerres, 2001). Wo liegen nun die Potenziale neuer Lernmedien? Darf man behaupten, dass das Lernen mit dem Computer effektiver wird, dass die Qualität des Gelernten steigt? Für einen sinnvollen Einsatz neuer Lerntechnologien ist die Überprüfung ihrer Wirksamkeit in realen Lernsituationen unabdingbar. Inwiefern computergestützte Lernmedien eine Bereicherung und eine wertvolle Ergänzung zum Unterricht sind, muss also sorgfältig evaluiert werden. Die Entwicklung neuer Lernsysteme wird auf der technischen Seite vorangetrieben. Eine wichtige, nicht zu vernachlässigende Aufgabe liegt jedoch in der genauen Überprüfung der Effekte und auch der Akzeptanz solcher Systeme bei den Benutzern sprich bei den Lernenden. Viele Evaluationsstudien zeigen ein ernüchterndes Ergebnis und können die Erwartungen in eine höhere Lernwirksamkeit computergestützter Medien gegenüber traditionellen Lernmedien nicht bestätigen. Die Forschungsbefunde hierzu sind aber auch widersprüchlich, was verallgemeinernde Aussagen über die Wirksamkeit computergestützter Medien wiederum nicht zulässt. Evaluationen als Mittel der Qualitätssicherung sind jedoch durchaus sinnvoll und notwendig, die Ergebnisse sollten aber nicht verallgemeinert, sondern lediglich auf das konkrete Lernsystem, seine Gestaltung und curriculare Einbindung interpretiert werden.

Im Mittelpunkt der vorliegenden Arbeit steht *Augmented Chemistry*, ein interaktives, computergestütztes Lernsystem, mit einer haptischen Benutzerschnittstelle (Abbildung 1.1). Mit *Augmented Chemistry* kann der Aufbau von einfachen Molekülen spielerisch erforscht werden. Das von vielen Lernenden als sehr theoretisch und abstrakt empfundene Fachgebiet der Organischen Chemie soll hiermit erfahrbarer, verständlicher und interessanter gemacht werden. Der Name dieses Lernsystems ist abgeleitet von *Augmented Reality* (erweiterte Realität), was soviel heisst

wie, dass die reale Welt mit virtuellen Elementen ergänzt wird. Augmented Chemistry wurde von Benedikt Vögtli (2002) als Diplomarbeit an der Fachhochschule beider Basel (FHBB), Abteilung HyperWerk, realisiert.

Augmented Chemistry wurde für eine Basisausbildung in Chemie konzipiert und dient zum Lernen von Molekülstrukturen. Das Lernprogramm soll einen leichteren Einstieg in die Chemie ermöglichen und Gelegenheit bieten, Freude am Thema zu entwickeln. Das Ziel der Entwickler ist es, Augmented Chemistry in einem realen schulischen Umfeld als Lernumgebung erfolgreich einzusetzen. Bevor dies geschehen kann, muss jedoch zuerst die Frage geklärt werden, wie effektiv Schüler mit diesem neuen System lernen, wie der Zugang der Lernenden zum Lernstoff ist, und ob dieser Zugang vielleicht sogar besser ist als mit herkömmlichen Lernmitteln. Dieser Frage soll in vorliegender Studie genauer nachgegangen werden. Da für die Brauchbarkeit eines Lernsystems nicht nur erfolgreiches Lernen wichtig ist, wurden in dieser Studie neben objektiven Leistungsdaten wie Lernerfolg auch subjektive Daten erhoben, welche indirekt Aufschluss über die Systemakzeptanz von Augmented Chemistry geben sollen. Subjektive Daten wurden anhand der Beanspruchung, der emotionalen Stimmung und der Usability oder Gebrauchstauglichkeit von Augmented Chemistry erhoben. Im Rahmen dieser Studie wurde die Lerneffektivität und Akzeptanz von Augmented Chemistry im Vergleich zu einem traditionellen Lernmittel evaluiert. Als Vergleichsmedium wurde das *Kugel-Stab-Modell* eingesetzt, ein Molekülbaukasten, womit reale Molekülmodelle anhand von Kugeln und Stäbchen zusammengebaut werden können (Abbildung 1.2).



Abb. 1.1: Lernsystem Augmented Chemistry.

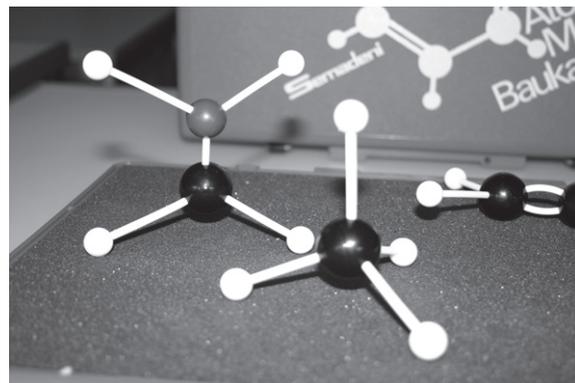


Abb. 1.2: Lernsystem Kugel-Stab-Modell.

Anhand dieser Vergleichsuntersuchung wurde analysiert, ob Personen, die mit dem virtuellen Lernsystem Augmented Chemistry lernen, einen vergleichbaren oder sogar besseren Lernerfolg aufweisen können, als Personen, die sich mit dem herkömmlichen nicht-computergestützten Lernsystem Kugel-Stab-Modell vorbereiten und ob es Akzeptanzunterschiede gibt. Die Unter-

schiede hinsichtlich der Lernwirksamkeit und der Akzeptanz der beiden Lernsysteme wurden anhand einer experimentellen Studie mit zwei Gruppen im Umfeld einer Berufsschule für Lehrlinge der chemischen und pharmazeutischen Industrie evaluiert. Eine Gruppe arbeitete dabei mit dem neuen System Augmented Chemistry, die andere Gruppe mit dem herkömmlichen Kugel-Stab-Modell. Die Effekte der Lernmethode auf die Lernwirksamkeit und die subjektiven Beurteilungen wurden durch wiederholte Messungen in einem AB/BA-Design untersucht, wobei jede Untersuchungsperson als seine eigene Kontrollperson wirkte. Für die Lernwirksamkeit wurden mit den Lernsystemen bearbeitete Aufgaben und ein Test ausgewertet. Die Beanspruchung wurde mit dem Nasa Task Load Index (Hart und Staveland, 1988) erhoben, einer mehrdimensionalen subjektiven Beanspruchungsmessmethode. Die Ermittlung der emotionalen Reaktion auf die Lernsysteme erfolgte mit dem Self-Assessment-Manikin (Lang, 1985), mit welchem Emotionen auf den drei Dimensionen Aktiviertheit, Valenz und Dominanz gemessen werden können. Eingesetzt wurde zudem ein eigens konstruierter Fragebogen zur Messung von subjektiven Urteilen auch hinsichtlich der Zufriedenheit mit der Lernsituation. Die subjektive Usability von Augmented Chemistry wurde schliesslich mit dem Software Usability Measurement Inventory (SUMI; Kirakowski und Corbett, 1993) erfasst.

Die Arbeit erhebt nicht den Anspruch, aus den Ergebnissen allgemeine Schlussfolgerungen für das Gebiet des Lernens mit Multimedia zu ziehen. Der Fokus liegt ausschliesslich auf Augmented Chemistry. Eine Theorie des multimedialen Lernens gibt es noch nicht, und empirische Untersuchungen für haptische Benutzerschnittstellen und Lernen fehlen weitgehend. Dies ist ein Grund, sich mit den Auswirkungen von Augmented Chemistry auf die Lernwirksamkeit und die Systemakzeptanz genauer auseinanderzusetzen. Bei der Evaluation von Augmented Chemistry wird man Eigenheiten finden und Optimierungsmöglichkeiten ausarbeiten. Elementare und generalisierbare Erkenntnisse werden aber rar bleiben.

Der Aufbau dieser Lizentiatsarbeit ist wie folgt: Im folgenden Kapitel 2 wird der theoretische Hintergrund erläutert. Dabei wird zuerst allgemein auf das Lernen mit computergestützten Lernsystemen eingegangen. Da jede Lernsoftware auf einer Lerntheorie basiert werden anschliessend kurz die drei Hauptströmungen Behaviorismus, Kognitivismus und Konstruktivismus erläutert. Anschliessend werden die Lernmedienforschung und ihre zentralen Ergebnisse behandelt. Nach einer kurzen Erläuterung des Computereinsatzes im Chemieunterricht wird das Lernsystem Augmented Chemistry genauer vorgestellt. Kapitel 3 enthält die Fragestellungen der Untersuchung. Kapitel 4 beinhaltet die Methode mit einer genaueren Beschreibung des Forschungsdesigns. Die Ergebnisse in Kapitel 5 gliedern sich nach den Fragestellungen. Eine Diskussion der Ergebnisse mit einer Bewertung der praktischen Konsequenzen der Studie erfolgt in Kapitel 6. Kapitel 7 gibt schliesslich eine kurze Zusammenfassung der wesentlichen Ergebnissen dieser Arbeit.

2. Theoretischer Hintergrund

Die folgenden Kapitel bilden den theoretischen Teil dieser Arbeit. Nach allgemeinen Überlegungen zum computerunterstützten Lernen wird genauer auf die drei psychologischen Lerntheorien des Behaviorismus, Kognitivismus und Konstruktivismus eingegangen. Es folgen weitere theoretische Überlegungen zur Messung von Lernwirksamkeit, Usability, Emotionen und Beanspruchung, um danach auf konkrete empirische Befunde zur Wirksamkeit computergestützter Lernsysteme einzugehen. Anschliessend werden die Einsatzmöglichkeiten des Computers als Lernmedium im Chemieunterricht erläutert. Zum Schluss werden die beiden Lernsysteme Augmented Chemistry und das Kugel-Stab-Modell genauer vorgestellt.

2.1 Lernen mit dem Computer

Im Folgenden wird nach einer kurzen Begriffsbestimmung für computergestützte Lernsysteme allgemein die Funktion von Medien im Bildungsbereich mit einem Blick auf die Mediendidaktik erläutert. Anschliessend wird auf die Nutzung neuerer Technologien zum Lernen eingegangen, im Speziellen auf das Prinzip von *Virtueller Realität*. In diesem Zusammenhang wird die Technologie von *Augmented Reality* ausführlicher beschrieben, da das zu untersuchende Lernsystem Augmented Chemistry auf dieser Technologie aufbaut.

2.1.1 Begriffsbestimmung

Multimedia, Telemedien, E-Learning, Computer Based Training (CBT), Computer Based Instruction (CBI), Computer Assisted Learning (CAL), Computer Assisted Instruction (CAI), Computerunterstützter Unterricht (CUU), Computerunterstütztes Lernen (CUL), Teachware, Courseware, Hypermedia, etc. Es gibt eine grosse Begriffsvielfalt für den Bereich computergestützter Lehr- und Lernsysteme. Eine genaue Abgrenzung dieser Begriffe ist schwierig und wird in der Regel gar nicht getroffen. Zudem umschreiben fast alle eingangs erwähnten Begriffe denselben Sachverhalt. Oft verwenden verschiedene Autoren auch einen dieser Begriffe als Oberbegriff für alle Arten von Lernsoftware (Mader & Stöckl, 1999). Im Folgenden wird versucht, etwas Ordnung in diesen Begriffswirrwarr zu bringen.

Eine erste grobe Unterteilung ergibt sich durch die Sichtweise. Je nachdem, ob der Aspekt des Lehrens oder Lernens im Mittelpunkt steht, wählt man Begriffe wie *Instruction/Unterricht* oder *Learning/Lernen* (vgl. Mader & Stöckl, 1999). Da in vorliegender Arbeit der *Lern-* und nicht der *Lehraspekt* im Vordergrund steht, und es sich um ein Lernsystem mit *lernunterstützender* Funktion handelt, wird der Begriff *computerunterstütztes Lernen* bzw. *computergestütztes Lernsystem* verwendet.

Es gibt eine grosse Vielfalt an Lernsoftware. Auf jeden einzelnen Typus von Lernsoftware einzugehen wird an dieser Stelle nicht beabsichtigt und würde auch den Rahmen dieser Arbeit sprengen. Es geht vielmehr darum, einen Überblick über die vorhandenen Softwaretypen zu schaffen. Eine Kategorisierung von Lernsoftware kann nach vielfältigen Kriterien erfolgen, unter anderem nach dem Grad der Interaktionsfreiheit des Lernenden gegenüber dem Programm (Schulmeister, 1997), nach dem Grad der Lernkontrolle, den die Lernprogramme zulassen (Ferguson, 1992; zit. nach Schulmeister, 1997, S. 66) oder sogar nach der jedem Lernsystem implizit zugrundeliegenden Lerntheorie (Blumstengel, 1998). Blumstengel (1998) wähnt zudem eine Typologie nach der grundlegenden Methodik als hilfreich. Methodische Grundtypen von Lernsoftware sind beispielsweise Hypertext-/Hypermedia-Lernsysteme, Tutorielle Systeme und Intelligente Tutorielle Systeme (ITS), Lernspiele, Simulationen, Drill- und Testsoftware, Mikrowelten oder Präsentations- und Visualisierungssoftware. Für genauere Erläuterungen zu diesen Lernsoftwaretypen sei auf die Autoren Schulmeister (1997), Blumstengel (1998) und Mader und Stöckl (1999) verwiesen.

Computergestützte Lernsysteme werden heutzutage auch oft mit dem Begriff *Multimedia* versehen. Multimedia ist in den letzten Jahren zu einem Schlagwort in der mediendidaktischen Diskussion geworden. Eine Definition erweist sich aufgrund mangelnden Konsenses in der wissenschaftlichen Diskussion als schwierig. Grundsätzlich geht es bei Multimedia aber immer um eine Integration verschiedener Informationsarten - von Texten, Grafiken, Ton und Film - in einem System (vgl. Euler, 1994, S. 292; Mader & Stöckl, 1999, S. 81; Kerres, 2001, S. 13). Es werden also gleichzeitig mehrere Wahrnehmungskanäle angesprochen. Schulmeister (1997) betont neben der integrativen Verwendung verschiedener Medientypen, dass bei Multimedia stets von einem interaktiven Medium die Rede ist. Eine wesentliche Erweiterung des Spektrums interaktiver Medien bietet der Bereich der *Virtuellen Realität*. Da das Konzept von Virtueller Realität, im Speziellen die Technologie von Augmented Reality, wesentlich zur Entwicklung von Augmented Chemistry beigetragen hat, wird dieses Thema in Kapitel 2.1.3 ausführlicher behandelt.

2.1.2 Computergestützte Lernsysteme und ihre Verwendung im Bildungsbereich

“Ein Lernmedium dient der Auseinandersetzung von Lernenden mit einem Lerngegenstand und beinhaltet Lernangebote, die das Erreichen bestimmter Lehrziele ermöglichen” (Kerres, 2001, S. 22). Bauer (1995) betrachtet die Interaktivität als entscheidendes Merkmal von multimedialen Lernprogrammen. Multimediales Lernen ermögliche ein entdeckendes und konstruierendes Lernen, unterstütze die Konstruktion von subjektivem Wissen, rege Gefühle an, und Simulationen würden zu einem tieferen Erleben führen (Bauer, 1995, S. 390). Interaktiver multime-

dialer Software wird das Potenzial zugesprochen, aktiv das Lernen zu unterstützen (vgl. Huk, 2003). Neben diesen vorwiegend positiv berichteten Eigenschaften von computergestützten oder multimedialen Lernmedien liegt aber auch eine eher skeptische Einstellung gegenüber diesem neuen Medium im Bildungsbereich vor, insbesondere bei den Lehrkräften.

Herrscht noch immer die Angst vor, Lehrer könnten durch die zunehmende Technisierung der Wissensvermittlung überflüssig werden? Bewahrheitet sich die Vision von Perelman (1992; zit. nach Schulmeister, 1997, S. 9), wonach Schulen gar geschlossen werden könnten? Vertreter dieser ökonomische Sichtweise erhoffen sich eine Kostenreduktion durch den Einsatz von medialen Lernprogrammen. Von dieser Extremposition ist man aber weggekommen, mehr hin zu einem integrativen Ansatz. Der Medieneinsatz wird vielmehr als Ergänzung zum Unterricht gesehen, das Medium versteht sich nicht mehr als Ersatz, sondern als Hilfsmittel für den Lehrer. Der Lehrer wird nicht überflüssig, er hat jedoch andere Aufgaben. Nicht mehr die Wissensvermittlung steht für ihn im Vordergrund, sondern die Unterstützung zum selbständigen Lernen durch Anregungen und Hilfestellungen (vgl. Bauer, 1995, S. 391). Der Lehrer wird zum Lernberater- oder -organisator und ist nicht mehr nur Wissensvermittler (Stangl, 2004). Eine veränderte Einstellung dem Lernen gegenüber ist zudem auf Seite der Lernenden unumgänglich; "Planungs- und Selbststeuerungskompetenz" werden vorausgesetzt (Strittmatter & Mauel, 1995, S. 48) und die Selbständigkeit und Eigenverantwortung der Lernenden müssen unterstützt werden. Nach Büchner (1995; zit. nach Bauer, 1995, S. 391) muss der Schüler beim Lernen mit dem Computer Lernstrategien entwickeln, die zum Erfolg führen, er muss seinen individuellen Weg finden, und er muss lernen, wie er am besten lernt. Daneben kann der Schüler aber frei über sein Arbeitstempo entscheiden, und der Computer informiert ihn über Mängel und Lernfortschritte.

Mediendidaktik

Mit der Frage, wie Medien als Hilfsmittel für einen erfolgreichen Unterricht eingesetzt werden können, beschäftigt sich die Mediendidaktik, die als ein Teilgebiet der Medienpädagogik verstanden wird: "In der Mediendidaktik geht es um Fragen der Entwicklung, Verwendung und Evaluation von Medien als Mittel zur Verbesserung von Lehr- und Lernprozessen." (Tulodziecki et al., 1996; zit. nach Blumstengel, 1998). Die Mediendidaktik befasst sich mit Fragen des didaktischen Gesamtverständnisses von Lernmedien und untersucht dabei, wie erfolgreicher Unterricht mit Hilfe von Medien zu bewerkstelligen ist (Strittmatter & Mauel, 1995). Die Funktion und Bedeutung von Medien in Lehr- und Lernprozessen ist ein zentrales Thema (vgl. Kerres 2004). Wichtige Funktionen des Lernmediums sind nach Kerres (2001):

- a) Wissens(re)präsentation: Darstellung und Organisation von Wissen
- b) Wissensvermittlung: Steuerung und Regelung des Lernprozesses
- c) Wissenswerkzeug: Konstruktion und Kommunikation von Wissen

In der Mediendidaktik geht es auch um die Nutzung von Medien in Lehr- und Lernprozessen. Das Ziel liegt in der Steigerung der Effektivität und Effizienz des Lehrens und Lernens durch den Einsatz von Technik (Kerres, 2004). Ein Lernmedium allein gilt aber noch nicht als Garantie für effektiveres Lernen. Nach Kerres und de Witt (2002) wird der Wert eines Mediums durch die Situation, nicht durch das Medium selbst bestimmt. Medien seien gar "Artefakte" ohne einen Wert an sich, Bedeutung erlangen sie erst durch eine bestimmte Nutzung in bestimmten Kontexten zu bestimmten Zeiten. Inwiefern der Einsatz computergestützter Lernsysteme wertvoll und gewinnbringend ist, hängt also stark vom Einsatzszenarium des Mediums ab. Der Einsatz computergestützter Medien ist nur unterstützend für das Lernen, wenn er "adäquat gewählt, sinnvoll und methodisch begründet" in den Unterricht eingebunden wird (Eilks, Krille & Flintjer, 2004). Der Lehrer als entscheidender Faktor für den Lernerfolg ist dabei nicht zu unterschätzen: entscheidend für eine wirksame Nutzung ist nämlich nicht nur das Programm selbst, sondern auch der Lehrer, der seine Schüler für die Arbeit mit dem Programm begeistern kann (Schulmeister, 1997). Ein nachhaltiger und wirkungsvoller Einsatz von computergestützten Medien in der Schule erfordert Kompetenzen sowohl im Umgang mit solchen Medien - wie sie eingesetzt werden - als auch in der kritischen Bewertung und Auswahl angebotener Medien (Eilks et al., 2004). Die Auseinandersetzung mit solchen Medien sollte demnach fester Bestandteil der Aus- und Weiterbildung von Lehrkräften sein (vgl. Klimsa, 1995, S.22).

Neben der mediendidaktischen Sicht verweist Kerres (2001) auf weitere Sichtweisen medien-gestützten Lernens. Dabei gibt es die technische, ökonomische, bildungsorganisatorische, medienwissenschaftliche und die lern- oder kognitionspsychologische Sicht. Letztere konzentriert sich vermehrt auf die Interaktion zwischen Lernendem und Medium im Lernprozess. Die "kognitive Funktion" von Medien beim Lernenden steht hierbei im Mittelpunkt (vgl. Kerres, 2001, S. 87).

Ein Grossteil mediendidaktischer Forschung ist der Frage nachgegangen, ob mediengestütztes Lernen erfolgreicher als konventioneller Unterricht ist, oder ob bestimmte neue Medien den alten Medien vorzuziehen sind. Werden im Unterricht computergestützte Lernsysteme eingesetzt, wird erwartet, dass ein näher zu spezifizierender Lernerfolg eintritt. Es stellt sich damit die Frage, ob bzw. wie das Lernen mit diesen Medien den Lernerfolg beeinflusst und worin der Mehrwert solcher Medien liegt. Es gibt eine Vielzahl von Untersuchungen, welche die Wirksamkeit von computergestützten Lernmedien evaluiert haben. Auf die Befunde dieser Evaluationsstudien wird in Kap. 2.3.5 genauer eingegangen.

2.1.3 Virtuelle Realität

Im Folgenden soll das Prinzip von Systemen, die auf Virtueller Realität basieren, genauer erläutert und das Potenzial solcher Systeme für den Bildungsbereich aufgezeigt werden.

Für Alsdorf und Bannwart (1995, S. 438) beginnt *Virtuelle Realität* (VR) dort, “wo ein virtuelles Modell durch die unmittelbare Reaktionszeit des Rechners auf die Manipulation des Benutzers hin interaktiv veränderbar wird.” Die Technologie entwickelte sich Ende der 1980er Jahre. Unter dem Begriff *Virtuelle Realität* wurden die damals aktuellen technischen Entwicklungen zur Schaffung “computergenerierter Erlebniswelten” und “interaktiver Mensch-Computer-Schnittstellen” zusammengefasst (Bente, Krämer & Petersen, 2002, S. 2). VR-Systeme sind immersive Computerwelten, die einem erlauben in eine gänzlich künstliche, respektive fiktive Welt, einzutauchen. Mittels spezieller Schnittstellen (z.B. Head-Mounted-Displays und Datenhandschuhen) wird dieses Eintauchen in die virtuelle Computerwelt ermöglicht und dem Nutzer das Gefühl vermittelt, von der dreidimensionalen Computerwelt umgeben zu sein (Alsdorf & Bannwart, 1995). Unter dem Begriff VR subsumiert sich eine Vielzahl unterschiedlicher Technologien. Eine an Biocca und Delaney (1995, S.59; zit. nach Bente et al., 2002, S. 10) angelehnte Klassifikation von VR-Systemen ist in Tabelle 2.1 wiedergegeben.

Tab. 2.1: Systemtypen, die auf Virtueller Realität basieren (nach Biocca und Delaney (1995)).

Typen von VR-Systemen	Beschreibung
Desktop VR-Systeme	Zur Darstellung einer virtuellen Welt wird lediglich ein konventioneller Computermonitor verwendet. Für stereoskopisches Sehen werden häufig 3D-Brillen eingesetzt.
Spiegel-Systeme	Der Nutzer betrachtet die Projektion eines Bildschirms und sieht ein Abbild seiner selbst, das sich in der projizierten virtuellen Welt bewegt.
Fahrzeug-basierte Systeme	Der Nutzer begibt sich hierbei in ein fahrzeugähnliches Vehikel (z.B. Flugzeug, Auto) und bedient die Kontrollen. Die Aussenwelt wird über einen Bildschirm abgebildet. In diese Kategorie gehören Flug- und Fahrsimulatoren.
Immersive VR-Systeme	Der Sichtpunkt des Benutzers wird komplett in das Innere einer virtuellen Welt verschoben. Diese VR-Systeme sind häufig mit Head-Mounted-Displays ausgerüstet.
Cave Systeme	Ein Cave ist ein Raum, an dessen Wänden virtuelle Bilder projiziert werden. Innerhalb dieses Raumes hat man das Gefühl sich in einer realen Welt zu befinden. 3D-Brillen werden verwendet, um den Raumeindruck zu verstärken.
Augmented Reality-Systeme	Die reale Welt wird mit virtuellen Elementen ergänzt. Computergenerierte 3D-Objekte können mit der Ansicht der realer Objektwelt überlagert werden. Halbdurchlässige LCD-Displays oder video-basierte Mischung von Umweltaufnahmen und Computer-Grafik-Output können hier eingesetzt werden.

Besonderes Augenmerk wird in vorliegender Arbeit auf die Technologie von *Augmented Reality* (AR) gelegt, da das zu untersuchende Lernsystem Augmented Chemistry auf dieser Technologie aufbaut.

Der Einsatz Virtueller Realitäten ist derzeit ein wichtiger Forschungsgegenstand (Huk, 2003). Die Nutzung dieser Technologie als Lernmedium wird in Zukunft an Bedeutung gewinnen. Charakteristisch für solche Lernumgebungen ist, "dass Lerner sich mit einer Lernumgebung konfrontiert sehen, die räumlich dreidimensional organisiert ist, die aktiv und selbstgesteuert erkundet werden kann, und deren visuelle Präsentation dynamisch dem Explorationsverhalten [...] angepasst ist." (Schwan & Buder, 2002, S. 109). Eine Lerntheorie in virtuellen Umgebungen ist stark an den Konstruktivismus angelehnt. Lernen durch Entdecken und Explorieren sind zentrale Annahmen (Winn, 1993). Virtuelle Lernwelten stellen zwar eine Bereicherung für den Lernenden dar, da sie die Möglichkeit haben, "eine virtuelle Welt zu erkunden, in ihr zu üben, zu experimentieren oder zu konstruieren" (Schwan & Bruder, S.119). Verschiedene Autoren versprechen sich von VR eine Steigerung des Interesses und der Motivation bei den Lernenden (vgl. Pantelidis, 1995; Winn, 1993). Inwiefern aber solche Lernwelten einen effektiven und effizienten Wissenserwerb zulassen, ist bislang eine weitgehend unbeantwortete empirische Frage. Die wenigen empirischen Studien zeigen zudem uneinheitliche Befunde. Es zeigte sich, dass insbesondere die Veranschaulichung abstrakter Konzepte (z.B. das Bindungsverhalten chemischer Moleküle oder das Verhalten elektrischer Felder) in virtuellen Lernwelten zu einem Lernzuwachs führten (Salzman, Dede, Loftin & Chen, 1999; zit. nach Schwan & Buder, 2002, S. 127). Ausschlaggebend sei neben der dreidimensionalen Veranschaulichung aber vielmehr die Möglichkeit zur interaktiven Manipulation gewesen. Kein bedeutender Lernzuwachs zeigte sich hingegen bei virtuellen Lernwelten, die gegenständliche Sachverhalte abbildeten (Bowman, Wineman & Hodges, 1999; Dede, Salzman & Loftin, 1996; Roussos et al., 1999; zit. nach Schwan & Buder, 2002, S. 127). Gegenständliche Sachverhalte sind möglichst realistische, illusionistische Darstellungen, beispielsweise Rekonstruktionen historischer Gebäude oder Fahrsimulatoren. Virtuelle Lernwelten haben demnach einen grossen Stellenwert in der Veranschaulichung abstrakter Sachverhalte. Sie bieten die Möglichkeit, Dinge zu erkunden, welche in der realen Welt nicht erkundbar sind. Schwan und Buder (2002, S. 127) sehen die Begründung der eher geringen Lerneffektivität von virtuellen Lernwelten darin, dass viele dieser Welten noch über relativ "informationsarme" Darstellungen verfügen. Verglichen dazu sind traditionelle medienbasierte Lernstoffpräsentationen teilweise viel differenzierter und bieten eine grössere Informationsdichte. Zukünftig müssen aber neben Fragen der technischen Realisierung auch pädagogisch-psychologische Perspektiven berücksichtigt werden. So kann die Abhängigkeit von Wissensvermittlung und Gestaltung virtueller Lernumgebungen vermehrt berücksichtigt werden.

Augmented Reality

Augmented Reality (AR) bezeichnet nach der Klassifikation in Tabelle 2.1 einen speziellen Typus von VR-Systemen. Während bei VR ein vollständiges Eintauchen in eine virtuelle Welt geschieht, wird bei AR die reale Welt nicht vollständig durch eine virtuelle Welt ersetzt. Die Realität wird lediglich mit virtuellen Elementen ergänzt. AR wird deshalb auch als erweiterte Realität bezeichnet. Bei AR kommt es zu einer Überlagerung der Ansicht realer Objektwelt mit computergenerierten 3D-Objekten (vgl. Bente et al., 2002). Einem AR-System kann man folgende Eigenschaften zuschreiben (Azuma et al., 2001, S. 34):

- AR kombiniert reale und virtuelle Objekte in einer realen Umgebung.
- AR arbeitet interaktiv und in Echtzeit.
- AR registriert reale und virtuelle Objekte und richtet diese aneinander/zueinander aus.

Bei AR-Systemen sind zwei verschiedene Darstellungsformen möglich. Es bietet sich die Möglichkeit, dass die reale Welt direkt betrachtet und die computergenerierte Information über halbdurchlässige Displays (z.B. Head-Up-Displays) zusätzlich eingeblendet wird. Solche Systeme werden beispielsweise in Flugzeugen eingesetzt. Der Pilot kann so gleichzeitig seine unmittelbare Umgebung und topographische Daten der Landschaft sehen (Abbildung 2.1).

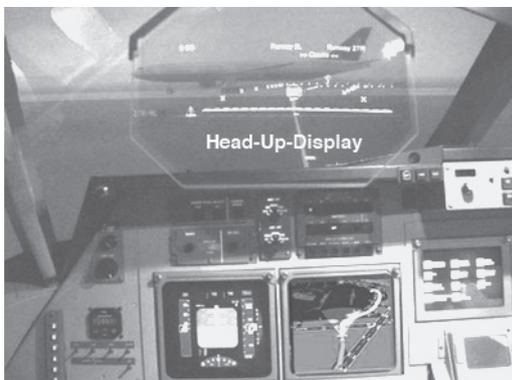


Abb. 2.1: Beispiel eines Head-up-Displays im Flugzeugcockpit.



Abb. 2.2: Beispiel eines Head-Mounted-Displays zur Darstellung einer Roboteranwendung.

Die andere Möglichkeit besteht darin, dass die Kombination von realer und virtueller Welt über ein nicht transparentes Display, ein Head-Mounted-Display, oder einfach über einen PC-Monitor betrachtet wird. Hierbei wird die Realität von einer Videokamera eingefangen und indirekt mit den zusätzlichen Informationen auf einen Bildschirm abgebildet (Abbildung 2.2). In diesem Fall ist es auch möglich, unerwünschte oder nicht relevante Objekte aus der Darstellung zu

entfernen (Zwisler, 1998). Die AR-Technologie ist nicht neu und findet in den verschiedensten Bereichen seine Anwendung. Das Potenzial von AR im Bildungsbereich wird aber erst allmählich untersucht. Der Vorteil des Einsatzes von AR als Lernmedium liegt darin, dass Lernende dreidimensionale Objekte auch dreidimensional sehen und wahrnehmen können. Basierend auf dem kollaborativen AR-System *Studierstube* (Schmalstieg et al., 2002), welches 1995 entwickelt wurde, entstanden unter laufenden Weiterentwicklungen des Systems Anwendungen unter anderem auch für den Bildungsbereich. Ein Beispiel einer solchen Anwendung ist Construct3D für den Geometrie- und Mathematikunterricht (Abbildung 2.3). Für einen Überblick über weitere Forschungsarbeiten zu *Studierstube* sei auf die Internetseite <http://www.studierstube.org/> hingewiesen.



Abb. 2.3: Die Arbeit mit Construct3D, einem kollaborativen AR-System.

2.2 Psychologische Lerntheorien

Baumgartner und Payr (1994) betrachten Lerntheorien im Zusammenhang mit computerunterstütztem Lernen als wichtig, da jede Lernsoftware auf einer Lerntheorie basiert, unabhängig davon ob dies explizit bei der Erstellung beabsichtigt wurde oder nicht. Im Laufe der Zeit entwickelten sich viele verschiedene Lerntheorien, die aber fast jede einer der drei Hauptströmungen Behaviorismus, Kognitivismus oder Konstruktivismus zugeordnet werden können. Im Folgenden sollen diese drei Theorien, insbesondere auch in Bezug auf das Lernen mit Computern, kurz skizziert werden.

2.2.1 Behaviorismus

Die ersten Konzepte zum computergestützten Lernen waren entscheidend durch die Theorien des Behaviorismus geprägt. Noch heute beeinflussen diese Modelle unsere Vorstellungen über mediengestütztes Lernen (Kerres, 2002). Im Behaviorismus sind interne Prozesse des Individuums nicht von Bedeutung. Das Verhalten wird lediglich durch Konsequenzen, die auf das Verhalten folgen, gesteuert. Ein grundlegender Mechanismus ist das Lernen durch Verstärkung, wonach erwünschtes Verhalten konditioniert werden kann, indem die Reaktion auf Reize verstärkt bzw. belohnt wird (Stimulus-Response-Modell). Erfolgreiches Verhalten wird eher gelernt als solches, das nicht mit Erfolg verbunden ist (Euler, 1994). Rückmeldungen sind ein wichtiges Prinzip im Lernprozess. Skinner versuchte das Stimulus-Response-Modell auf das Lernen am Computer zu übertragen und legte damit in den 1960er Jahren den Grundstein für die ersten Konzepte des Computereinsatzes im Bildungsbereich. Die Anwendung dieser Konzepte wird als Programmierter Unterricht bezeichnet. Lerninhalte werden dabei in kleinste Schritte unterteilt und in eine festgelegte Reihenfolge gebracht. "Das Programm arbeitet damit eine vorprogrammierte Sequenz von Lernschritten ab." (Kerres, 2002, S. 58). Die Lernschritte sollten leicht zu bewältigen sein und mit Prüfungen abgeschlossen werden. Durch die ständige Bekräftigung soll hierbei ein stetiger Lernerfolg eintreten. Dem Lernen am Computer wurde eine hohe Behaltensleistung sowie Transferleistung des Gelernten zugeschrieben. Die Hoffnung lag darin, Lehrende überflüssig zu machen, da der Computer als der geeignetere Informationsvermittler angesehen wurde (Stangl, 2004).

Zusammengefasst liegt der lerntheoretischen Richtung des Behaviorismus "das passive rezeptive Lernen in kleinsten Schritten mit jeweiligem Feedback" zugrunde, und "die Vermittlung, die Instruktion, das Lehren und die deduktive Vorgehensweise" stehen im Vordergrund (Issing, 1995, S. 197). Gelernt wird hier lediglich Faktenwissen. Das Verstehen, Anwenden oder Bewerten komplexer Zusammenhänge wird ausgeschlossen (Euler, 1994). Das stereotype und einseitige Lernen im Programmierten Unterricht konnte sich nicht durchsetzen. In sogenannten Drill- und Testsoftware, womit bereits gelernte Inhalte gefestigt und kognitive Fertigkeiten eingeübt werden können (z.B. Vokabellernen), finden die Konzepte des Behaviorismus aber nach wie vor ihre Berechtigung.

2.2.2 Kognitivismus

Die kognitive Psychologie geht auf die Theorien von Jean Piaget und Jerome S. Bruner zurück. Im Zentrum der Forschung steht der Mensch als informationsverarbeitendes System. Er wird dabei als ein "aktives und selbststeuerndes System" betrachtet (Euler, 1994, S. 296). Es gibt eine Vielfalt von Ansätzen, aber alle haben gemeinsam, dass der menschliche Denkprozess

als ein Prozess der Informationsverarbeitung angesehen wird (Mader & Stöckl, 1999). Kennzeichnend hierfür ist die durch den Kognitivismus geprägte *Computer-Metapher*: Informationen werden dabei, ähnlich wie bei einem Computer, als Input aufgenommen, intern verarbeitet und letztlich als Output wieder ausgegeben.

Im Kognitivismus wird Lernen als besonderer Fall der Informationsaufnahme und –speicherung betrachtet und geht einher mit Veränderungen kognitiver Strukturen und Prozesse (Bonner, 1998; Shuell, 1987, zit. nach Kerres, 2001, S. 66). Lernen wird als ein Prozess der Aneignung von Wissen im Gedächtnis eines Individuums aufgefasst. Dieses Wissen ist in der Person gespeichert und kann später aus dem Gedächtnis abgerufen bzw. rekonstruiert werden (Kerres 2001, S. 75). Der Lernprozess wird also von innen gesteuert und ist abhängig von den jeweils vorhandenen subjektiven Erfahrungsstrukturen (Euler, 1994).

Aus der kognitiven Psychologie heraus hat sich das Modell des *Entdeckenden Lernens* entwickelt (vgl. Schulmeister, 1997, S. 71). Hierbei ist - in Bezug auf das Lernen mit dem Computer - die Freiheit des Lernenden wichtig, Wege und Strategien im Umgang mit Programmen selbst wählen zu können. Es wurde das Ziel verfolgt, adaptive Systeme zu entwickeln, die sich besser auf den Lernfortschritt der Lernenden einstellen (Kerres, 2001). Hierzu zählen die Tutoriellen Systeme.

Spätestens seit Ende der 1980er Jahre hat am kognitiven Ansatz eine grundlegende Kritik eingesetzt. Menschliches Handeln würde auf kognitive Informationsverarbeitung reduziert, bei der das Individuum als Zentrum von Wissen und Handeln überbewertet und “die menschliche Emotionalität, Leiblichkeit und Situiertheit des Handelns in der Lebenswelt” ausgeblendet wird (Kerres, 2001, S. 74). Computeranalogie und Affektlosigkeit, Ausklammerung von Situation, Kultur und individuellem Aspekt sind also die Hauptkritikpunkte (Mader & Stöckl, 1999).

Neue Ansätze geben das Konzept der kognitiven Informationsverarbeitung nicht grundsätzlich auf, vielmehr bewerten sie deren Bedeutung anders.

2.2.3 Konstruktivismus

Als Alternative zu kognitivistischen Ansätzen werden zunehmend konstruktivistische Ansätze des didaktischen Designs diskutiert.

Der Grundgedanke des Konstruktivismus liegt in der „Genese des Wissens“ (Schulmeister, 1997, S. 73). Die Theorie befasst sich mit der Formation und Bedeutung des Wissens. Wissen wird demnach konstruiert und existiert nicht unabhängig vom Subjekt. Das heisst, dass jedes Individuum seine eigene Wirklichkeit konstruiert. Es gibt kein objektives Wissen. Lebewesen sind nach Maturana und Varela (1987) informationell geschlossene Systeme, d.h. sie nehmen

Informationen nicht wie objektive Gegebenheiten auf, sondern nur nach eigenen Regeln durch Interpretation (vgl. Mader & Stöckl, S. 40; Schulmeister, 1997, S. 77;). Wissen lässt sich folglich nicht durch Instruktion vermitteln, sondern muss vom Lernenden aktiv in seine “mentalen Modelle und Wirklichkeitskonstrukte” integriert werden (Issing, 1995, S. 197). Zu einer Erweiterung des Wissens kommt es durch Anbindung neuer Informationen an bereits bestehendes Wissen. Es geschieht eine „Koppelung“ von Vorwissen und aktueller Situation (Mader & Stöckl, 1999, S. 41).

Lernen ist aus konstruktivistischer Sicht ein Konstruktionsprozess. Der Lernende setzt sich aktiv mit Informationen auseinander, und dabei steht er selbst und nicht der Lernstoff im Mittelpunkt. Die Eigenständigkeit und Flexibilität der Lernenden wird betont; richtiges Lernen ist autonomes und aktives Lernen. Der Lehrer wird eher als Coach oder Lernhelfer betrachtet. Die Verantwortung für den Lernerfolg liegt somit beim Lernenden selbst (Mader & Stöckl, 1999). Die Autonomie des Individuums muss also stärker beachtet werden. Aus pädagogischer Sicht bedeutet dies, dass Lernumgebungen entwickelt werden müssen, in denen kognitive Lernprozesse in handelnder Auseinandersetzung mit der Umwelt stattfinden können (Schulmeister, 1997). Offene Lernumgebungen sollen “authentische Situationen in ihrer Komplexität und realitätsnah präsentieren bzw. Interaktionen des Lerners mit diesen Situationen simulieren [...]” (Strittmatter & Mauel, 1995, S. 55).

Welche Rolle wird den Medien in konstruktivistischen Ansätzen zugeschrieben? Medien werden nicht als „Behälter“ verstanden, in denen Wissen gespeichert ist und übermittelt werden soll, Medien sind vielmehr „Werkzeuge“, um Wissen zu konstruieren (Kerres, 2001, S. 82). Medien werden deshalb auch als “kognitive Werkzeuge” für den Lernenden bezeichnet (vgl. Jonassen, 1991; zit. nach Issing, 1995, S. 198). Dieser Werkzeugcharakter soll die aktive Wissenskonstruktion unterstützen. Es gibt noch wenige Forschungsergebnisse zum Einsatz von konstruktivistischen Lernmodellen, vielleicht gerade deshalb, weil diese Lernmethoden nur schwer mit herkömmlichen Kontrollverfahren überprüfbar sind (Huk, 2003).

2.3 Wirksamkeit computergestützter Lernsysteme

In den folgenden Kapiteln wird zuerst allgemein auf das Konzept der Evaluation eingegangen um anschliessend die Messung von Lernerfolg in Evaluationsstudien genauer zu betrachten. Nach einer Erläuterung verschiedener Aspekte, welche bei Evaluationen berücksichtigt werden sollten, wie die Usability und die Beanspruchung, wird konkret auf empirische Befunde zur Wirksamkeit computergestützter Lernsystem eingegangen. Der Abschluss bilden Überlegungen zu Einschränkungen von Evaluationsstudien.

2.3.1 Evaluationsforschung im Bildungsbereich

Baumgartner (1999) macht auf die Definitionsproblematik, Vielschichtigkeit und Komplexität von Evaluation aufmerksam und zeigt verschiedene theoretische Ansätze und Verfahrensweisen auf. Es soll nicht Aufgabe dieser Arbeit sein, all diese Ansätze zu beleuchten und voneinander abzugrenzen. Es wird deshalb hier nur auf einzelne Aspekte eingegangen. Baumgartner (1999) hebt für den Evaluationsbegriff insbesondere den Prozess der Bewertung hervor.

Der Begriff Evaluation im neueren deutschen Sprachgebrauch ist aus dem Englischen übernommen und stammt von *to evaluate*, was soviel bedeutet wie bewerten, begutachten oder (zahlenmässig) bestimmen oder berechnen (Grüner, 1993). Evaluationsdefinitionen im Zusammenhang mit Bildung heben folgende Punkte hervor: Zielorientiertheit, Systematik, Kontinuität, Bewertung und Objektorientierung (Grüner, 1993; Will, Winteler & Krapp, 1987, zit. nach Fricke, 1995, S. 405). Grüner (1993, S. 31) definiert Evaluation in Bezug auf Bildungsmaßnahmen demnach folgendermassen: „Evaluation ist der zielgerichtete systematische Prozess der Informationssuche und Bewertung dieser Information, bezogen auf ein ausgewähltes Bildungsprogramm.“. Gegenstand der empirischen, pädagogisch-didaktischen Wirkungsforschung ist die Wirkung spezieller Inhalte, Lernformen, didaktischer Konzepte oder Lehrstile auf den Lernerfolg, auf die Zufriedenheit und auf die Motivation von Lernenden mit speziellen Bildungsangeboten (Schulmeister, Vollmers, Gücker & Nuyken, 2003). An dieses Grundmodell der empirischen Evaluationsforschung ist auch das Evaluationsdesign von Lernsoftware angelehnt.

Es gibt eine Vielzahl von Varianten der Evaluation. Je nach Schwerpunkt der Zielsetzung und Realisationsmöglichkeiten haben sich unterschiedliche Evaluationskonzepte, sogenannte Evaluationsmodelle entwickelt (Wottawa & Thierau, 1990). Angesichts der zahlreichen unterschiedlichen Modellansätze hat für die Projektgestaltung vor allem Scrivens Unterscheidung zwischen formativer und summativer Evaluation die grösste Bedeutung (Wottawa & Thierau, 1990). Die formative Evaluation findet während der Entwicklungsphase eines Programms statt und dient der Optimierung und Beseitigung von Schwachstellen eines Systems. Die summative Evaluation dagegen kommt erst bei vollständig entwickelten Programmen zum Einsatz und dient der Bewertung und Beurteilung der Effektivität eines Programmdesigns. Die summative Evaluation legt ihren Fokus auf die Qualität eines bereits durchgeführten Programms. Das Ergebnis, der Erfolg oder der Nutzen von Massnahmen sind hierbei von Interesse (vgl. Blumstengel, 1998; Fricke, 1995; Grüner, 1993; Issing, 1995; Wottawa und Thierau, 1990).

Im Folgenden wird das Evaluationskonzept CIELT (Concept for Interdisciplinary Evaluation of Learning Technologies) von Grund, Windlinger und Grote (2002) kurz erläutert (siehe Abbildung 2.4). Dieses Evaluationskonzept kann als Leitfaden dienen für eine gezielte Durchführung

einer Evaluation. Aussagen auf bestimmten Evaluationsebenen (in der Abbildung ganz rechts) sind an Voraussetzungen bezüglich der Verwendung des Systems durch die Zielgruppe gebunden. Die gegenseitige Abhängigkeit dieser Voraussetzungen werden im Modell vereinfacht durch eine Pyramide veranschaulicht. Die Ausgangspunkte sind Systemstabilität und Zugänglichkeit über verschiedene Browser gekoppelt mit einem technischen Support, der im ersten Schritt die grundsätzliche Nutzung des Systems gewährleistet. Darauf baut die Akzeptanz des Lernmediums und der curricularen Einbindung auf, woraus sich ein empirisch beobachtbares Nutzerverhalten ergibt. Anschliessend findet idealerweise eine nach dem didaktischen und methodischen Konzept des Entwicklers realisierte Systemnutzung statt, wonach spezifische Lerneffekte untersucht werden können. Abschliessend ist die Dauer des Einsatzes und die Nutzung der Technologie Voraussetzung dafür, dass sich Veränderungen auf organisationaler Ebene untersuchen lassen. Sämtliche gewonnenen Erkenntnisse fliessen zurück in die Systemkonzeption und das Curriculum. Nach diesem Modell werden in vorliegender Studie lediglich Usability und Lernfacetten untersucht. Die Voraussetzungen dafür werden als gegeben betrachtet. Die gewonnenen Erkenntnisse sollen wieder ins didaktische Konzept, ins Design und die curriculare Integration zurückfliessen.

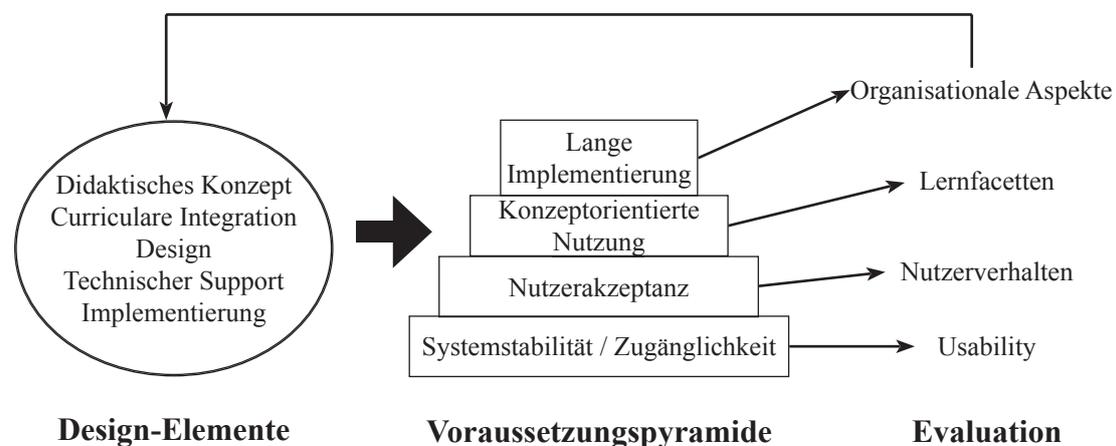


Abb. 2.4: CIELT-Rahmenmodell für die Evaluation von Lernmedien (aus Grund et al., 2002).

Die pädagogische Evaluationsforschung ist am Design des psychologischen Experiments mit Versuchs- und Kontrollgruppe orientiert. Als unabhängige Variablen werden in den Versuchsbedingungen die Lernformen eingeführt. Als abhängige Variablen werden individuelle Lernerfolge sowie Einstellungen, Denkstile, Motivationslagen und Beurteilungen der Lernsituationen durch die Lerner erfasst. Der Lernerfolg wird im Anschluss an die Lernphase meist als Prüfungsleistung in Form von Punkten oder Noten gemessen (Schulmeister et al., 2003).

2.3.2 Messung des Lernerfolgs von computergestützten Lernsystemen

Die Erfassung von Lernerfolg ist mit gewissen Schwierigkeiten verbunden. Denn was unter Lernerfolg verstanden wird und dementsprechend die Operationalisierbarkeit dieses Konstrukts, ist ein kontrovers diskutiertes Thema. Spricht man von Lernerfolg, wenn etwas schneller, erfolgreicher oder nachhaltiger gelernt wird?

Untersucht man Lernerfolg, ist wahrscheinlich von Interesse, ob etwas gelernt wird, und wenn ja wie gut. Es interessiert ob und wie die Lernziele erreicht wurden. Glowalla und Schoop (1992; zit. nach Glowalla & Häfele, 1995, S. 427) haben folgende Kriterien zur Beurteilung der Wirksamkeit computergestützter Lernmedien vorgeschlagen: "Man muss erfassen, wieviel Wissen erworben wird, wieviel Zeit ein Lerner zum Erwerb dieses Wissens benötigt und welchen Lernweg er einschlägt. Darüber hinaus muss erfasst werden, wie der Lernende die Arbeit mit dem jeweiligen System einschätzt". Schliesslich sollten zur Qualitätssicherung alternative Kursangebote in kontrollierten Umgebungen miteinander verglichen werden. Eine Möglichkeit bildet der parallele Einsatz einer Kontrollgruppe, bei der eine andere Unterrichtsmethode angewendet wird. Um die Gleichwertigkeit der Referenzmethode sicherzustellen, muss die zum Vergleich herangezogene Massnahme aber in bestmöglicher Qualität angeboten werden.

In Evaluationsstudien wird Lernerfolg weitgehend mit Behaltensleistung gleichgesetzt (Kerres, 2001). Zur Kontrolle des Lernerfolgs werden in den meisten Fällen Prüfungen angewendet, die deklaratives und prozedurales Faktenwissen testen. Diese Sichtweise wird insofern problematisch, da zunehmend ein konstruktivistisches Lernverständnis einsetzt, wodurch die eigene *Wissenskonstruktion* nicht die *Wissensreproduktion* in den Vordergrund rückt. Nach Dubs (1993; zit. nach Euler, 1994, S. 299) darf die Evaluation des Lernerfolgs demnach nicht auf Lernprodukte mit ausschliesslich richtigen und falschen Lösungen ausgerichtet werden, sondern zu überprüfen seien die Fortschritte bei den Lernprozessen. Auch Baumgartner (1999, S. 93) macht auf die Bedeutung von Lernprozessen aufmerksam: so "müssen bei der Evaluation interaktiver Medien vor allem die durch die Interaktivität angestossenen Lernprozesse evaluiert werden". Seine Überlegung gründet darin, dass Schüler mit interaktiven Lernprogrammen den Lerngegenstand erforschen und sich in konstruktiver Weise Wissen aneignen können. Benutzer sind daher nicht mehr bloss "Rezipienten", sondern können in den "medial vermittelten Informations-, Kommunikations- und Lernprozess gestaltend eingreifen" (Baumgartner, 1999, S. 90). Herkömmliche Methoden für die Messung des Lernerfolgs wie die Reproduktion von Wissensinhalten sei nicht mehr adäquat für die Beurteilung von Lernerfolg durch den Umgang mit interaktiven Medien. Die Lösung liegt folglich in der Beobachtung und Bewertung von Lernprozessen (vgl. auch Kozma, 1994; zit. nach Freudenreich & Schulte, 2002, S. 2). Das Protokollieren von Lernprozessen ist heute mit Hilfe geeigneter Software möglich. Dafür

werden speziell entwickelte Programme eingesetzt, welche automatisch jegliche Benutzeraktivität im System protokollieren und die Informationen in sogenannten *log files* speichern. Log file-Untersuchungen haben den Vorteil, dass mit relativ geringem Aufwand eine Vielzahl von objektiven Daten erhoben werden kann, die sich durch eine gute Vergleichbarkeit auszeichnen (vgl. Degenhardt, 2001).

2.3.3 Usability

Unter *Usability* versteht man ganz allgemein die Benutzbarkeit eines Produkts, eine hypothetische Eigenschaft, die dem Produkt zugeschrieben wird, wenn es beispielsweise “benutzerfreundlich, angenehm zu bedienen, geeignet zum Erfüllen einer bestimmten Aufgabe” ist (Hegner, 2003, S. 32). Übersetzt wird der Begriff Usability auch häufig mit Gebrauchstauglichkeit oder Brauchbarkeit. Eine zufriedenstellende Benutzbarkeit ist für den Erfolg von Lernsoftware äusserst wichtig. Nach Boling und Sousa (1993; zit. nach Schaumburg & Rittmann, 2000) stellt die Benutzbarkeit eine entscheidende Voraussetzung dar, dass ein Lernsystem überhaupt zum Lernen genutzt wird. Laut Bevan (1999) können zuverlässige Messungen der Gebrauchstauglichkeit nur erlangt werden, wenn repräsentative Benutzer auch realistische Aufgaben in realistischen Umgebungen bearbeitet haben. In jüngster Zeit haben sich Evaluationsmethoden etabliert, die eine Beurteilung der Software durch die Benutzer selbst anstreben. Standardisierte Fragebögen sollen die Messung einzelner, genau definierter Usability-Kriterien und damit die Vergleichbarkeit zwischen unterschiedlichen Produkten ermöglichen.

Es gibt internationale Standards, die sogenannten ISO-Normen, welche Qualitätsmerkmale von Usability beschreiben. In ISO 9241-11 wird hervorgehoben, dass Usability vom Nutzungskontext abhängt und dass die besonderen Umstände, unter denen das Produkt genutzt wird, den Grad an Usability bestimmen. Usability wird demnach definiert als das Ausmass, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und mit Zufriedenheit zu erreichen (Abbildung 2.5).

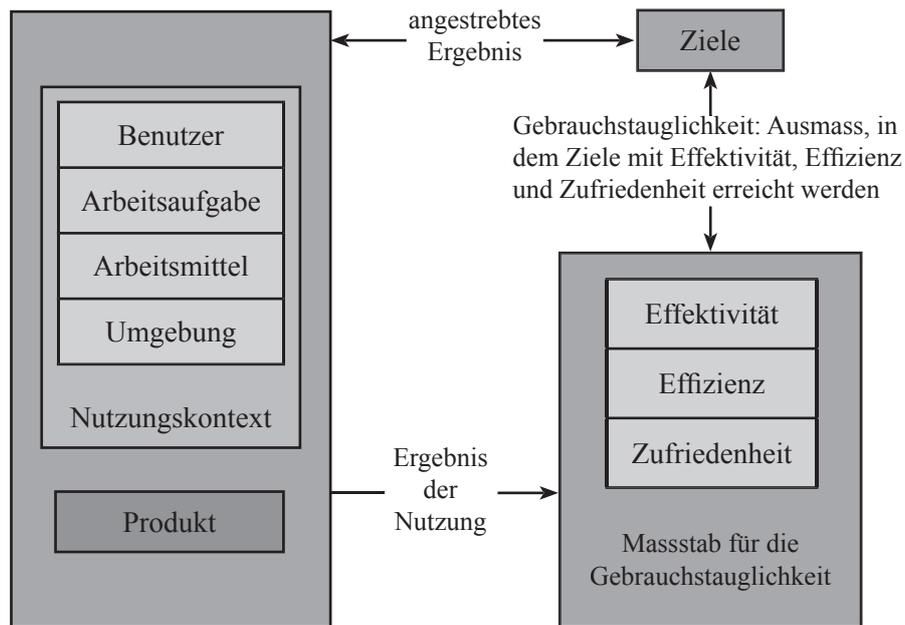


Abb. 2.5: Anwendungsrahmen der Gebrauchstauglichkeit nach ISO-Norm 9241-11.

Nach diesem Modell werden die drei Usability-Komponenten wie folgt definiert:

- Effektivität: Die Genauigkeit und Vollständigkeit, mit der Benutzer ein bestimmtes Ziel erreichen.
- Effizienz: Der im Verhältnis zur Genauigkeit und Vollständigkeit eingesetzte Aufwand, mit dem Benutzer ein bestimmtes Ziel erreichen.
- Zufriedenheit: Beeinträchtigungsfreiheit und positive Einstellung gegenüber der Nutzung des Produkts.

Diese mit Usability eng in Verbindung gebrachten Begriffe der Effektivität und Effizienz eines Softwaresystems deuten darauf hin, dass die Gebrauchstauglichkeit eines solchen Systems stark von funktionsgebundenen Kriterien abhängig gemacht wird (Trautmann, 2004). Gebrauchstauglich ist eine Software demnach erst dann, wenn sie eine ausreichende Funktionalität besitzt und ihre Benutzung dem Nutzer auch vermittelt werden kann. Die Zufriedenheit der Benutzer stellt sich als Konsequenz ein und ist wiederum für die Akzeptanz des Produkts verantwortlich. Der Zufriedenheitsbegriff nach ISO 9241-11 stellt zwar eine subjektive Komponente im Bewertungsprozess dar, ist aber inhaltlich ebenfalls sehr aufgaben- oder funktionsgebunden. Die Zufriedenheit ist auch definiert als subjektive Einstellung der Benutzer gegenüber dem Produkt und ist sehr bedeutend, da das subjektive Empfinden des Benutzers oft entscheidend für die Akzeptanz eines Systems ist (Hegner, 2003). Neben der Lernwirksamkeit ist die Akzeptanz von Lernsystemen Schwerpunkt vieler Evaluationsstudien, da sie unter anderem auch zum Lernerfolg beiträgt (vgl. Blumstengel, 1998; Kerres, 2001).

Die klassische Perspektive einer funktionsgebundenen Systembewertung wie sie unter der ISO-Norm 9241-11 erfolgt, wird zunehmend kritisiert. Die Forderung nach einem erweiterten Konzept von Usability mit einer eingehenderen Beschäftigung mit subjektiven Wahrnehmungen und Empfindungen macht sich breit (Dillon, 2001; Hassenzahl, Beu & Burmester, 2001). Bis anhin wurde die Zufriedenheit als einzige affektive Komponente von Usability berücksichtigt. Dillon (2001) bemängelt dies und fordert, sich eingehender mit subjektiven Aspekten wie dem Erleben von Software zu beschäftigen. Emotionen als Usability-Masse sollen nicht weiter vernachlässigt werden. “However, even the most rigid rationalist must agree that much of what we do as humans is driven by emotional and affective factors” (Dillon, 2001, S. 61). Die Rede ist auch von *Joy-of-Use* (vgl. Hassenzahl et al., 2001). Im Bereich Mensch-Maschine-Interaktion ist Joy-of-Use eine Erweiterung des Usability Engineering durch eine emotionale Komponente. Man spricht deshalb auch von *Emotional Usability*. Eine der am besten untersuchten affektiven Konzepte im Zusammenhang mit der Mensch-Maschine-Interaktion ist die wahrgenommene Freude bei der Benutzung. Freude trägt deshalb zur Akzeptanz eines Programms bei, da sie als intrinsischer Motivator gilt (Davies et al., 1992; zit. nach Trautmann, 2004, S. 23). Der wahrgenommenen Freude wird sogar eine positive Auswirkung auf die Computerangst zugeschrieben (Igarria, Schiffman & Wieckowski, 1994; zit. nach Trautmann, 2004, S. 23). Dem Zusammenwirken von System und Emotion sollte also in zukünftigen Evaluationsstudien vermehrt Rechnung getragen werden. Das Nutzererleben und deren emotionale Komponente in der Systembewertung rücken damit immer mehr ins Zentrum der Aufmerksamkeit.

In vorliegender Arbeit wird auf die ISO-Norm 9241-11 hinsichtlich der beiden Usability-Komponenten Effektivität und Zufriedenheit Bezug genommen. Diese beiden Usability-Komponenten werden durch die Lernerfolgsmessung, bzw. durch die subjektiv empfundene Usability von Augmented Chemistry, erfasst. Um subjektiven Aspekten wie dem Erleben von Software ebenfalls Rechnung zu tragen, wird auch die emotionale Reaktion auf Augmented Chemistry erhoben.

2.3.4 Messung der Beanspruchung

Neben der Lernleistung ist es sinnvoll, die Beanspruchung durch eine Lernsituation zu erfassen. Eine Ergänzung der Leistungsbewertung durch eine Analyse der mit einer Aufgabe verbundenen Beanspruchung ist in der ergonomischen Forschung von grosser Bedeutung (vgl. Hacker & Richter, 1986; Ulich, 1994). Beanspruchungsanalysen sind auch bei der vergleichenden Bewertung von alternativen Mensch-Maschine-Systemen einzusetzen, da sich Systeme hinsichtlich aufzubringendem Aufwand für die Leistung stark unterscheiden können (Pfundler, 1990). So können zwei verschiedene Lernsysteme zwar gleich gute Lernleistungen bei den Lernenden hervorrufen, aber die Lernenden sehr unterschiedlich in Anspruch nehmen, was sich wieder-

um auf die subjektive Beurteilung und Akzeptanz der Lernsysteme auswirken kann. Mentale Beanspruchung wirkt sich also nicht immer auf den Lernerfolg aus, dafür hat sie aber einen Einfluss auf die Zufriedenheit der Nutzer, welche neben der Leistung eine wichtige Voraussetzung für die Brauchbarkeit eines Systems darstellt (siehe Kapitel 2.3.3). Daneben kann eine zu geringe oder zu hohe Beanspruchung zu Langeweile oder Aufmerksamkeitsmangel führen, was wiederum die Effektivität des Arbeitens negativ beeinflusst (Bevan, 1995). Die mentale Beanspruchung hat hier einen Einfluss auf den Lernerfolg, was sich auch auf die Brauchbarkeit eines Lernsystems auswirkt. Als wichtigste Klasse der Beanspruchungsmessmethoden gelten die subjektiven Verfahren.

2.3.5 Empirische Befunde zur Wirksamkeit computergestützter Lernsysteme

Bei der Erprobung neu entwickelter Lernsysteme nimmt deren Überprüfung und Evaluation einen zentralen Stellenwert ein. Bereits aus den 1960er und 1970er Jahren liegt eine Vielzahl empirischer Untersuchungen zur Effektivität des Einsatzes von didaktischen Medien vor. Insbesondere zum computerunterstützten Lernen finden sich zahlreiche Ergebnisse aus Evaluationsstudien. Aufgrund der Fülle von wissenschaftlicher Literatur zu Fragen des computerunterstützten bzw. multimedialen Lernens soll vorgängig auf einige Metaanalysen eingegangen werden. Kulik und Kulik (1991) haben mehrere Metaanalysen zum Vergleich traditioneller Formen computer- und videogestützter Lernprogramme durchgeführt. Sie untersuchten die Effektstärke von mehreren hundert Einzelstudien zu computerunterstütztem Lernen (CUL) und fanden im Schnitt eine Effektstärke von 0.35, was ein eher kleiner Effekt ist. Bei der Lerndauer ergab sich eine Reduktion um 30% mit CUL gegenüber konventionellem Unterricht. Im Bereich der Einstellung zum Lernstoff beträgt die Effektstärke 0.25. Kulik (1994; zit. nach Blumstengel, 1998), der mehrere Metaanalysen analysierte, kommt zu folgenden Schlussfolgerungen: Insgesamt ist eine Verbesserung des Lernerfolgs durch CUL festzustellen, die Lernzeit kann relativ deutlich reduziert werden, CUL wird von den Lernenden überwiegend positiv aufgenommen, insgesamt ist aber kaum eine Verbesserung der Einstellung zum gelernten Fachinhalt nachweisbar. Fricke (1991; zit. nach Kerres, 2001, S. 106, Glowala & Häfele, 1995, S. 422) macht aber auf Einschränkungen der Aussagekraft von Metastudien aufmerksam, welche Ergebnisse einer grösseren Zahl einzelner Evaluationsstudien zusammenfassend auswerten: Die einzelnen Studien betreffen oft sehr unterschiedliche Systeme und verwenden keine einheitliche Evaluationsmethodik. Oftmals werden in den einzelnen Studien auch verschiedene Massnahmen vermischt oder der Lernerfolg wurde nicht sinnvoll erhoben oder nur geschätzt (vgl. Joliceur & Berger, 1986; zit. nach Hasebrook, 1995, S.95). Kulik und Kulik (1991) machen auf ein weiteres Problem aufmerksam: In Metastudien fehlen aktuelle Befunde von Evaluations-

studien über die Wirksamkeit computergestützter Lernsysteme. Da Computer einem enormen Wandel in kürzester Zeit unterlegen sind - sie werden kleiner, schneller, billiger und zuverlässiger - müssen frühere Metastudien aktualisiert werden. Bereits eine Übersichtsstudie von Kulik und Kulik von 1987 fand Hinweise auf einen zeitlichen Trend in Berichten zur Effektivität von computergestützten Lernsystemen.

Gemäss Kerres (2001) zeigen zahlreiche Studien, dass die Effektivität von computergestützten gegenüber traditionellen Medien oder gegenüber konventionellem Unterricht nicht höher ist. Motivationale Vorzüge neuer Medien sind vielfach auf Neuigkeitseffekte zurückzuführen und für den einzelnen Benutzer nicht von Dauer. Ein Indiz für die Auswirkung des Neuigkeitseffekts ist die Beobachtung, dass der Lernzuwachs sich verringert, je vertrauter die Lernenden mit dem neuen Medium werden (vgl. Schulmeister, 1997). Besondere Vorteile des mediengestützten Lernens zeigen sich meist bei Personen, bei denen die Lernmotivation sehr hoch ist, und welche auch ein selbständiges Lernverhalten aufweisen. Die Akzeptanz von Lernangeboten in realen Lernsituationen und ihre Bedeutung für den Lernerfolg wurden bisher kaum untersucht, obwohl die Akzeptanz eine essentielle Bedingung für den Lernerfolg ist (Kerres, 2001).

Im Bereich der Naturwissenschaften gibt es eine Vielzahl von Untersuchungen zur Effektivität computerunterstützten Lernens. Einige Studien ergaben, dass computerunterstütztes Lernen effektiv in der Steigerung der Leistung von Schülern und ihrer Einstellungen gegenüber den Naturwissenschaften war (Chang, 2000; Ferguson & Chapman, 1993; Yalcinalp et al., 1995; zit. nach Chang, 2001). Auf der anderen Seite fanden Forscher, dass konventionelle Lernmedien gleich gut oder sogar besser waren, als das computerunterstützte Lernen (Morell, 1992; Wainwright, 1989; zit. nach Chang, 2001). Hasebrook (1995) verweist auf existierende Studien und Übersichtsarbeiten, die zwar gezeigt haben, dass Multimediasysteme über Potenziale zur Verbesserung der Lernleistung verfügen, die überwiegende Mehrzahl der heute im Einsatz befindlichen Multimediasysteme dennoch nur wenig oder gar keine positive Auswirkung auf die Lernleistung hat. Schulmeister (1997) stellt zudem fest, dass die meisten experimentellen Vergleiche von Unterrichtsmethoden keine signifikanten Resultate erzielen, und die wenigen signifikanten Befunde sich noch gegenseitig widersprechen. Es herrscht also grosse Uneinigkeit über die Effektivität computergestützter Lernsysteme im Unterricht. Den Gründen für solch entgegengesetzte Befunde und auf allgemeine Einschränkungen von Evaluationen wird in Kapitel 2.3.6 genauer eingegangen.

2.3.6 Methodische und praktische Einschränkungen von Evaluationen

Ein Grund für die widersprüchlichen Befunde in der Evaluationsforschung liegt zum einen darin, dass verschiedene Lernmedien schwer zu vergleichen sind, und deshalb allgemeine Aussagen zum Effekt solcher Medien kaum möglich sind. Für die Unmöglichkeit der Generalisierung von Aussagen über das einzelne untersuchte Produkt hinaus macht Schulmeister (1997) den geringen Grad an Typisierungsfähigkeit verantwortlich, der aus der breiten Varianz im Design und in den Lernmethoden resultiert. Auch Hasebrook (1995) weist darauf hin, dass es nicht möglich ist, prinzipielle Aussagen über die Lernwirkungen von Multimedia zu machen. Zudem ist unklar, welche Programmtypen in den einzelnen Untersuchungen, die in Übersichtsarbeiten zusammengefasst sind, evaluiert wurden. Welche Programmtypen werden unter dem Begriff computergestützte Lernsysteme zusammengefasst? Kann man von einem getesteten Produkt überhaupt auf andere schliessen? Es ist also festzuhalten, dass aufgrund praktischer und methodischer Einschränkungen Evaluationen nur begrenzt Hinweise auf die grundsätzliche Lernwirksamkeit von computerunterstütztem Lernen liefern. Evaluationen sind als Mittel der Qualitätssicherung aber dennoch sinnvoll und notwendig. Die Ergebnisse sollten aber lediglich auf das konkrete Lernsystem, seine Gestaltung und curriculare Einbindung interpretiert werden (Blumstengel, 1998).

Neben der geringen Generalisierungsfähigkeit von evaluativen Einzelstudien liegt ein weiteres Problem darin, dass sich das computerunterstützte Lernen in einem multidisziplinären Umfeld mit unterschiedlich starkem Praxisbezug befindet, was einen geringen Konsens über anzuwendende Forschungsstandards zur Folge hat (vgl. Grund, Windlinger & Grote, 2002). Summative Evaluationen tragen zu wenig mediendidaktisch relevanten Verlaufs- und Ergebnisvariablen, wie der erlebten Qualität des Lernangebots, des Lernverhaltens, den emotionalen Reaktionen, der Lernmotivation und der faktischen Nutzung mediengestützter Lernangebote Rechnung (vgl. Kerres, 2000; zit. nach Grund et al., S. 176). Zudem beläuft sich in vielen Untersuchungen die Nutzungsdauer von neuen Lernmedien nur auf 30 Minuten (Aufenanger, 1999), was keine differenzierte Aussagen über Lern- und Nutzerverhalten erlaubt, geschweige denn Aussagen über den Lernerfolg.

Evaluationsstudien stehen vor dem grossen Problem einer unendlichen Vielzahl von Lernervariablen gegenüberzustehen, die alle zu kontrollieren ein unmögliches Unterfangen darstellt. Versucht man den Einfluss der Lernervariablen durch zufällige Verteilung der Lernenden auf verschiedene Lernsituationen konstant zu halten, werden durch die Vielfalt der Lernstile, Einstellungen und Erwartungen unter den Lernenden die Unterschiede in der Wirkung der unabhängigen Variablen wiederum ausgeglichen (vgl. Schulmeister et al., 2003). Der amerikanische Psychologe Thomas L. Russell hat Studien analysiert, die computerunterstütztes Lernen mit

traditionellem Unterricht verglichen haben. Er fasst seine Befunde unter dem Titel “No Significant Difference Phenomenon” (Russell, 1999; zit. nach Schulmeister et al., 2003) zusammen. Wie dem Titel bereits zu entnehmen ist, konnte Russell aufzeigen, dass deutlich mehr nicht-signifikante Studien vorliegen als signifikante. Er kommt somit zum Schluss, dass sich empirisch nicht belegen lässt, ob sich eine bestimmte Unterrichtsmethode grundsätzlich als anderen überlegen erweist, wenn nicht spezifische Lernvoraussetzungen und Einstellungen mitberücksichtigt werden. Allein Lernsituationen miteinander zu vergleichen ist nicht sinnvoll, da derartige Studien zwangsläufig zu nicht-signifikanten Ergebnissen führen würden (Carey, 2001). Sobald man aber verschiedene Lernsituationen hinsichtlich didaktischer Aspekte (z.B. Ausmass der Einwirkungsmöglichkeiten durch den Nutzer) voneinander abgrenzt und die diesen Lernsituationen ausgesetzten Personen in Bezug auf spezifische Lernvariablen (z.B. extrinsisch oder intrinsisch motiviert, viel oder wenig Vorwissen) in Gruppen einteilt, kann man zu pädagogisch interessanten und statistisch signifikanten Ergebnissen gelangen.

2.4 Chemie und Computer

Die folgenden Ausführungen stützen sich auf Aussagen in einer Broschüre der Arbeitsgruppe “Computer im Chemieunterricht” der Fachgruppe Chemieunterricht in der Gesellschaft Deutscher Chemiker (Eilks et al., 2004). In der Broschüre “Computer und Multimedia im Chemieunterricht heute” wird der Versuch unternommen, in Zeiten sich stark verändernder Ansprüche an Schule, Lernen und Bildung den Aspekt des Einsatzes von Computer und Multimedia im Chemieunterricht aus didaktischer und lerntheoretischer Sicht kritisch zu betrachten. In einer über 20 Jahre dauernden Diskussion um das Lernen mit dem Computer im Chemieunterricht hat man sich viel von diesem neuen Medium für das Lernen versprochen. Viele aber nicht alle dieser Versprechen konnten eingelöst werden. Eine rasante Entwicklung der Technik in den letzten Jahren hat einen grundlegenden Wandel ermöglicht. Diese Entwicklung hat hin zu Lernprogrammen mit völlig neuen Perspektiven geführt, insbesondere durch den Einbezug verschiedener Medien (Multimedialität) und Darstellungsformen (Multicodalität) oder dem Ansprechen verschiedener Sinne (Multimodalität). In Bezug auf lernpsychologische Mechanismen bieten sich Chancen, das Lernen mit multimedialen Lernumgebungen auch im Sinne des konstruktivistischen und situativen Lernens zu legitimieren. Es wurde aber noch nicht abschliessend untersucht, ob und wie diese grundlegenden Theorien auf das Lernen mit den neuen Medien angewandt werden können.

Computersoftware zum Lernen von Chemie gibt es seit den ersten Personalcomputern in den 1980er Jahren. Der Wert für ein effektives Lernen und die unterrichtliche Verwertbarkeit für solche Lernsoftware ist nur für beispielhafte Fälle untersucht worden. Die uneinheitlichen Er-

gebnisse machen generelle Einschätzungen aber schwierig (Sumfleth & Kummer, 2001). Die Nutzung des Computers im Chemieunterricht ist nach wie vor kein zentrales Arbeits- und Forschungsfeld von Hochschulen. Hierin besteht also noch erheblicher Forschungsbedarf. Die vielfach lediglich unterstellten Effekte des Computereinsatzes im Chemieunterricht sind empirisch zu hinterfragen und gegebenenfalls zu untermauern. Insgesamt muss man feststellen, dass zurzeit nur sehr wenige Angebote an computergestützten Lernsystemen vorhanden sind, die nachgewiesen auf die Bedürfnisse schulischen Lernens von Chemie angepasst sind. Zudem sind kaum konzeptionelle Vorschläge vorhanden, wie die Softwareangebote im Chemieunterricht eingesetzt werden können. Folgende Aspekte bedürfen aus der Sicht der Fachdidaktik an Untersuchung:

- Computerbasierte Lernprogramme sollen verstärkt bestimmte Schritte des schulischen Lernens von Chemie unterstützen. Die Gestaltung sollte sich vermehrt am Lernprozess der Schülerinnen und Schüler und den Curricula orientieren
- Eine Evaluation der Einsetzbarkeit des entsprechenden Mediums im Unterricht muss bei der Erstellung der Medien berücksichtigt werden.
- Aus Sicht der Chemiedidaktik muss eine Untersuchung der benutzten Medien hinsichtlich ihrer lernunterstützenden Wirkung stattfinden. Das Ziel liegt darin, einzuschätzen mit welchem Umfang, welcher Komplexität oder in welchem unterrichtsmethodischen Zusammenhang ein solches Medium sinnvoll eingesetzt werden kann. Aufgrund unterschiedlicher Curricula verschiedener Schulen sind oft nur Fallstudien, aber keine allgemeine Beurteilung möglich.

Was Lernprogramme über Molekülbau betrifft, so besitzen animierte räumliche Darstellungen von Molekülen für mechanistische Betrachtungen ein beachtenswertes didaktisches Potenzial. Anwendungsbereiche sind beispielsweise die Vorstellung vom räumlichen Aufbau der Moleküle, Feststoff- und Festkörperstrukturen und Isomerie, aber auch die Visualisierung von Reaktionsmechanismen des Protonenaustauschs (Säure-Base-Reaktionen). Die didaktischen Konsequenzen wurden bisher allerdings weder auf der Ebene der universitären Ausbildung noch für die Schule tiefer gehend empirisch untersucht. Eine Diskussion des Potenzials der Moleküldarstellung mit dem Computer in Bezug auf das Modelldenken geben u.a. Reiners und Saborowski (2001).

Das, was man zumindest bei kleinen Molekülen mit den entsprechenden Molekülbaukästen in der Schule machen kann, lässt sich ergänzend auch am Computer durchführen. Im Folgenden werden exemplarisch einige Molekülbauprogramme vorgestellt. *Rasmol* ist das wohl am meisten verbreitete Molekülvisualisierungsprogramm. Mit Rasmol können plastische 3D-Darstellungen von Molekülen erstellt werden, und es ermöglicht die Visualisierung von Prote-

inen, Nukleinsäuren und kleinen Molekülen. *Ghemical* ist ein weiteres Programm, mit welchem Moleküle modelliert und berechnet werden können. Neben üblichen Funktionen wie Rotation, Translation, Zoomen, Löschen und Hinzufügen von Bindungen und verschiedenen Darstellungsmöglichkeiten ist eine Ansicht für 3D-Brillen eingebaut. Ein graphisch orientiertes Molekülbauprogramm ist *ChemSite*. ChemSite mit dem Untertitel "Der Molekülbaukasten für den Chemieunterricht" ist ein vielseitig verwendbares Werkzeug, mit welchem nicht nur beliebig grosse (organische) Moleküle konstruiert werden können. Man wird zudem in die Lage versetzt, an den Molekülen umfangreiche Messungen durchzuführen, die Moleküldynamik bei beliebigen Temperaturen zu simulieren und das Verhalten des Moleküls unter dem Einfluss von anderen Substanzen zu erforschen. Auf der Homepage von CTI (Center for Chemistry; <http://www.liv.ac.uk/ctichem/intro.html>) findet sich eine Datenbank welche über 500 Einträge von Software enthält, welche im Chemieunterricht verwendet werden kann. Welche Anzahl reine Molekülvisualisierungsprogramme ausmachen, wurde nicht gezählt. Es wird aber ebenfalls eine beachtliche Anzahl sein. Es ist also sehr viel auf dem Markt vorhanden, und diese Molekülbauprogramme sind teilweise sehr vielseitig verwendbare Werkzeuge. Dadurch werden sie aber insbesondere für Anfänger in der Organischen Chemie sehr schnell zu komplex und unübersichtlich.

Ein mit Augmented Chemistry vergleichbares Molekülbauprogramm gibt es nicht auf dem Markt. Im Vergleich zu den üblichen Benutzerschnittstellen, bei welchen über Maus und Tastatur mit dem Computer interagiert wird, soll AC erfahrbarer, intuitiver und einfacher zu bedienen sein.

2.5 Das Lernsystem Augmented Chemistry

Augmented Chemistry (AC) ist ein interaktives, computergestütztes Lernsystem, mit dem der Aufbau und die Struktur von einfachen Molekülen gelernt werden kann. AC richtet sich in erster Linie an eine Zielgruppe von Lernenden im Bereich der Organischen Chemie, die über keine oder nur geringe Kenntnisse in diesem Gebiet verfügen. Dazu zählen Lehrlinge der chemischen oder pharmazeutischen Industrie. AC soll das von vielen Lernenden als sehr theoretisch und abstrakt empfundene Fach der Organischen Chemie näher bringen, einen leichteren Zugang zu diesem Fach ermöglichen, sowie mehr Spass am Lernen vermitteln. Das Lernen von komplexen dreidimensionalen Molekülstrukturen erfordert ein gutes räumliches Vorstellungsvermögen. Solche abstrakten Lerninhalte sollen mit AC durch die dreidimensionale Darstellung besser veranschaulicht und somit verständlicher gemacht werden. Man verspricht sich von AC motiviertere Schüler, grössere Behaltensleistung, verbesserte Qualifikationen im beruflichen Umfeld und stärkere Identifikation des Lernenden mit dem gelehrteten Fach.

In den folgenden Kapiteln wird das Lernsystem AC genauer vorgestellt. Nach einem kurzen Bericht über die Entstehungsgeschichte von AC wird auf das System selber, sowie deren Funktionsweise und Benutzung eingegangen. Danach werden die didaktischen Motive von AC erläutert.

2.5.1 Entstehungsgeschichte

AC ist bei HyperWerk FHBB (Fachhochschule beider Basel) als Diplomarbeit von Benedikt Vögli entstanden, welcher die Idee und das Konzept zu dieser Anwendung nach einer intensiven Auseinandersetzung mit neuen Benutzerschnittstellen und interaktiven Lernsystemen entwickelte. Es zeichnete sich eine rasante Entwicklung ab: Zusammen mit Patrick Juchli, der für die Software-Architektur verantwortlich war, konnte das geplante Lernsystem innerhalb von nur zwei Monaten bis im März 2002 umgesetzt werden. Um die Kohärenz der chemischen Regeln sicherzustellen, wurden Fachleute vom Institut für Physikalische Chemie der Universität Basel beigezogen. Bei der Entwicklung von AC war der didaktische Wert der Anwendung für Benutzer ohne Vorbildung in der Molekularchemie von entscheidender Bedeutung. Für die Anwendung von AC sind nur geringe Computerkenntnisse erforderlich.

Im März 2002 wurde AC bei der Diplomausstellung bei Hyperwerk ausgestellt, wo ein erster Kontakt mit Graviton GmbH, eine in Basel tätige Software-Entwicklungsfirma mit Schwerpunkt E-Learning, zustande kam. Bald entstand eine Zusammenarbeit mit Graviton. Bei Graviton wurde innerhalb von 6 Monaten der Grundstein gelegt für einen erfolgreichen Einsatz von AC an Schulen und wissenschaftlichen Museen. Schliesslich wurde AC an den Cybernarium Days 2003 in Darmstadt (<http://www.cybernarium.de>), einer Ausstellung über virtuelle Welten, einem grossen Publikum gezeigt.

Bei der Entwicklung der Diplomarbeit von Benedikt Vögli hat auch Dr.sc.techn. Morten Fjeld von der Forschungsgruppe Mensch-Maschinen-Interaktion (MMI) an der ETH Zürich unterstützend mitgewirkt. Seine Initiative und seine hohe Fachkompetenz auf dem Gebiet der Human-Computer-Interaction haben das Projekt entscheidend vorangetrieben.

2.5.2 Systembeschreibung

Die Entwickler benutzten einen Ansatz aus dem Bereich Augmented Reality (siehe Kapitel 2.1.3) und kombinierten diese Technologie mit einem *Tangible User Interface* (TUI). *Tangible* heisst übersetzt greifbar, fühlbar. Ein TUI ist eine haptische Benutzerschnittstelle, bei welcher Objekte aus dem alltäglichen Umfeld des Menschen als Schnittstelle zum Computer verwendet werden. Ziel dieser Schnittstelle ist es, die reale mit der virtuellen Welt zu verbinden. Im Ge-

gensatz zum üblichen Graphical User Interface (GUI), bei dem der Benutzer über eine Maus oder die Tastatur mit dem Computer interagiert und Objekte somit nur indirekt manipulieren kann, ermöglicht TUI eine direkte Interaktion zwischen Mensch und Computer. Der Computer selbst rückt dabei in den Hintergrund. Die Kommunikation mit dem Computer soll dabei einfacher und vor allem intuitiver werden.

Bei AC können die Benutzer durch die Kombination von Augmented Reality System und TUI also mit virtuellen Molekülen interagieren, indem sie reale Objekte als Eingabewerkzeuge manipulieren. Diese realen Objekte ersetzen die Computertastatur und wurden speziell für die Arbeit mit dreidimensionalen Modellen entworfen. Bei den Objekten handelt es sich um einen Würfel, einen Greifer, ein Buch und verschiedene Karten, welche im Folgenden genauer beschrieben werden.



Buch

Jede Buchseite repräsentiert ein Element aus dem Periodensystem. (Grösse der Buchseiten: A5)



Greifer

Damit können Elemente aus dem Buch erfasst werden, um das Molekülmodell zu bauen.



Würfel

Damit kann das Molekülmodell um seinen eigenen Mittelpunkt gedreht, von allen Seiten betrachtet und in die richtige Position gebracht werden, um neu hinzukommende Atome einzubinden. (Material: Styropor; Kantenlänge ca. 8 cm)



Plattform

Darauf befindet sich das virtuelle Molekülmodell. Sie kann frei um die vertikale Achse gedreht und in alle Richtungen bewegt werden.



Karten

Insgesamt vier verschiedene Kärtchen dienen der Interaktion mit dem System: Library-Funktion zum Betrachten fertiger Moleküle, Darstellungsform der Moleküle wechseln, Plattform löschen, Ringstrukturen bilden. Es können auch Elektronegativität und Polarität dargestellt werden.

AC bietet sowohl eine dreidimensionale Darstellung von virtuellen Molekülen innerhalb der realen Welt, als auch eine besonders intuitive Form der Interaktion mit diesen Molekülen (Fjeld, Juchli & Vögli, 2003). Zusammengefasst sind folgende Aspekte von AC wesentlich:

- Die Interaktion mit virtuellen Objekten erfolgt mit Hilfe von realen Objekten.
- Der Benutzer kann beide Hände gleichzeitig zur Interaktion verwenden.
- Mehrere Objekte können gleichzeitig manipuliert werden.
- Eine Zusammenarbeit von mehreren Benutzern ist möglich.

Abbildung 2.6 veranschaulicht die Arbeitsfläche von AC. Der Arbeitsplatz von AC besteht aus einem Tisch und einer Rückprojektionswand, auf der die gewünschten Moleküle erscheinen. Unter der Projektionswand befindet sich eine Kamera, die jede Aktion vor dem Bildschirm erfasst. Das von der Kamera aufgenommene Bild wird dann über einen Beamer von hinten auf die Projektionswand projiziert, was dem Benutzer den Eindruck eines Spiegels verleiht (*mirror metaphor*). Das Spiegelbild kann dann mit virtuellen Elementen erweitert werden (Fjeld et al., 2003). Auf diese Art visualisiert AC dreidimensionale Atom- und Molekülmodelle.

Die Eingabewerkzeuge Würfel, Buch, Greifer und Karten sind mit spezifischen Mustern versehen, die von der Kamera aufgenommen und vom Computer erkannt werden. Je nachdem, was der Benutzer mit den Gegenständen macht, wird die Position der Muster verändert. Der Computer erkennt so die Veränderungen und steuert passend dazu das Bild auf der Projektionsfläche.

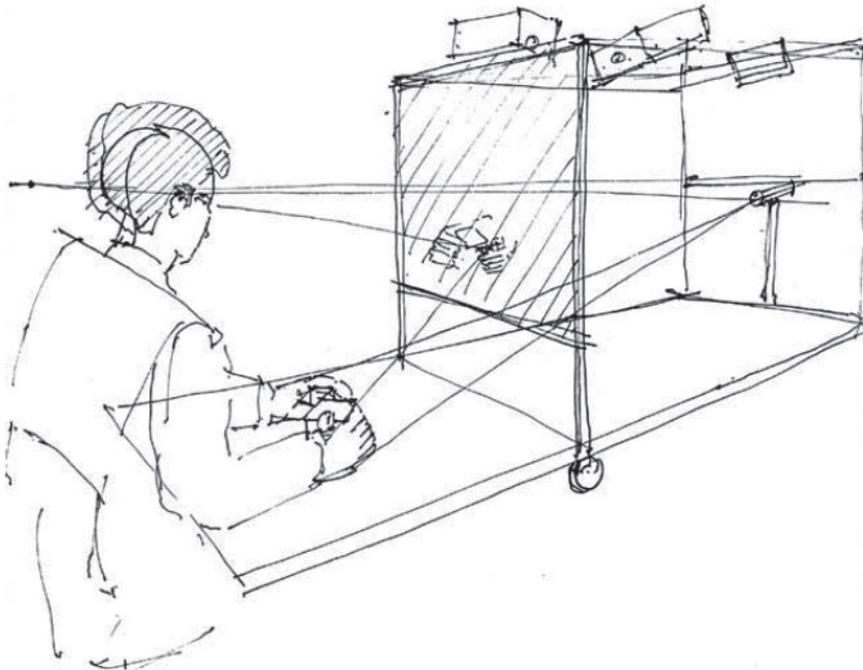


Abb. 2.6: Systemskizze von AC (im realisierten System befindet sich die Kamera jedoch näher beim Benutzer).

2.5.3 Funktionsweise und Benutzung

Um Molekülmodelle mit AC zu konstruieren, wählt man zuerst aus dem Buch ein Element aus (Abbildung 2.7). Mit dem Greifer erfasst man dieses Element, indem man ihn zum Buch hinbewegt (Abbildung 2.8). Mit einem Tastendruck auf den Greifer setzt man das Atom schliesslich auf die Plattform (Abbildung 2.9).



Abb. 2.7: Funktionsschritt 1.



Abb. 2.8: Funktionsschritt 2.



Abb. 2.9: Funktionsschritt 3.

Mit dem Würfel wählt man dann aus, wo eine Bindung zusätzlicher Atome an das Molekülmodell platziert wird. Je nach Lage des Würfels kann auch bestimmt werden, ob man eine Einfach-, Zweifach- oder Dreifachbindung wünscht. Zur Bestimmung der Bindungsart ist ein Tetraeder um das Zielatom aufgespannt. Ecke, Kante oder Fläche stehen für Einfach-, Zweifach-, oder Dreifachbindung (Abbildung 2.10).

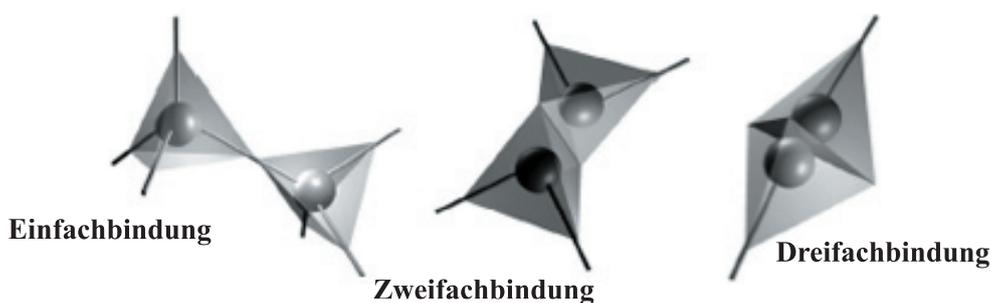


Abb. 2.10: Symbolisierung der Bindungsarten bei AC durch das Tetraedermodell.

Sobald ein Molekül gesättigt ist, gibt das System ein positives Feedback aufgrund eines akustischen Signals. Zusätzlich werden spezifische gesprochene Lerninhalte über das gebaute Molekül (z.B. Aggregatzustand, Vorkommen, gefährliche Eigenschaften, etc.) dargeboten. Diese auditiv präsentierten Informationen ertönen jeweils automatisch, nachdem das entsprechende Molekül gesättigt ist.

2.5.4 Didaktische Motive

In einem Computerlernprogramm sind bestimmte didaktische Konzepte vergegenständlicht: Art der Präsentation des Stoffs, Ansprache des Lerners, Ausmass und Form der Interaktivität, sowie Möglichkeiten des Eingriffs des Lernenden in den Ablauf der Software.

Mit AC lassen sich Molekülmodelle nach dem Baukastenprinzip zusammensetzen, was beispielsweise Chemielehrlinge beim Lernen von Molekülen unterstützen soll. Es geht insbesondere darum, abstrakte und komplexe Darstellungsformen der Moleküle anschaulicher und erfahrbarer zu machen. Ein exploratives Lernen soll hierbei unterstützt werden. Bei den Lerninhalten handelt es sich um Grundlagen der Organischen Chemie im Bereich Molekülbau und Molekülstruktur. In Tabelle 2.2 sind die Prinzipien aus der klassischen Strukturlehre abgebildet, welche mit Hilfe von AC vermittelt werden können. Neben diesen Prinzipien können zusätzlich durch auditiv präsentierte Lerntexte spezifische Eigenschaften der gebauten Moleküle gelernt werden.

Tab. 2.2: Prinzipien der klassischen Strukturlehre, die mit AC erarbeitet werden können.

Prinzip	Inhalt
Oktettregel (Edelgasregel)	Moleküle streben danach, acht AussenElektronen zu haben, da dies zu einer besonders stabilen Struktur führt. Somit versucht jedes Atom durch die Bindung mit anderen Atomen, die Anzahl seiner AussenElektronen auf acht Stück zu bringen. Die Anzahl Valenzelektronen eines Atoms gibt an, wieviele Bindungen dieses Atom eingehen kann.
Isomerie	Isomere sind Moleküle, welche dieselbe Summenformel haben, aber sich durch die Art, wie die einzelnen Atome im Molekül verbunden sind, unterscheiden.
Mehrfachbindungen, Ringe	Mehrwertige Atome können Einfach-, Doppel- oder Dreifachbindungen eingehen und sind oft als lineare oder verzweigte Ketten oder als Ringe miteinander verbunden.
Molekülgeometrie	Doppel- und Dreifachbindungen sind kürzer und ergeben eine andere Geometrie als Einfachbindungen. Liganden ordnen sich räumlich so an, dass die gegenseitige Abstossung der Valenzelektronen minimal ist. Für Atome mit vier Einfachbindungen ist somit die günstigste Anordnung diejenige eines Tetraeders. Geometrie um Mehrfachbindungen: Spannt man um jedes Atom ein Tetraeder, so haben einfach gebunden Atome eine Ecke, zweifach gebundene Atome eine Kante und dreifach gebundene Atome eine Fläche des Tetraeders gemeinsam.
Molekulares Dipolmoment	Verschiedene Atome ziehen Bindungsatome in verschiedenem Masse an. Das molekulare Dipolmoment ergibt sich aus der Vektorsumme der einzelnen Bindungsdipolmomente.

Der Einsatz von computergestützten Lernsystemen ist für verschiedene Arten von Lernzielen unterschiedlich gut geeignet. Unterschieden werden folgende Lernziele (vgl. Schanda, 1995): kognitive Lernziele, affektive Lernziele und psychomotorische Lernziele. Bei AC steht Ersteres, der unterstützende Erwerb kognitiver Fähigkeiten im Vordergrund. Bei den kognitiven Lernzielen steht die Vermittlung von Wissen im Sinne von Faktenwissen bzw. Wissen um Zusammenhänge im Vordergrund. Bei AC könnte das inhaltliche Ziel beispielsweise darin bestehen, dass der Lernende in der Lage sein soll, bestimmte Moleküle unter Berücksichtigung der Oktettregel zu bauen, die Moleküle zu benennen und ihre Strukturformel zeichnen zu können. Des Weiteren sollten Lernende spezifische Eigenschaften der Moleküle kennen. Beim Erwerb der dazu notwendigen Kompetenzen sind sowohl reines Auswendiglernen als auch einsichtiges Lernen beteiligt (Schanda, 1995).

Bei AC sind keine Lerneinheiten implementiert. Um AC in den Unterricht zu integrieren und damit einen Lernerfolg zu erzielen, ist es notwendig passend zu den zu unterrichtenden Lerninhalten Lernaufgaben zu entwickeln, die dann mit AC gelöst werden sollen. Die Lernenden er-

halten schriftliche Aufgaben, welche sie während einer Lektion einzeln bearbeiten müssen. Die Aufgaben müssen so gestellt sein, dass die Lernenden während der Bearbeitung etwas Neues dazulernen. Dieses selbständige Erarbeiten neuen Wissens ist die Kurzform des Entdeckenden Lernens (Frey, 2002). Das Modell des Entdeckenden Lernens geht auf kognitionspsychologische Ideen von Bruner (1961; zit. nach Schulmeister et al., 2003, S. 2) zurück. Bruner liess Schüler selbstständig Lösungen entdecken und unterstützte somit ihre Problemlösefähigkeit. Dieses eigene Entdecken ist gemäss Bruner ein Lernprozess, in dem Lernende ihr Wissen durch eigene Aktivitäten aufbauen. Entdecken bezieht sich somit auf konkrete physische Aktivitäten und begleitende kognitive Prozesse. Die Selbstkonstruktion von Wissen durch aktives Ausprobieren hat insbesondere Jean Piaget in der kognitiven Entwicklung untersucht (Schulmeister et al. 2003) und geht einher mit dem konstruktivistischen Lernverständnis.

In Bezug auf die psychologischen Lerntheorien lässt sich AC sicherlich im konstruktivistischen Paradigma einordnen (siehe Kap. 2.2). AC soll nämlich die Funktion eines Werkzeugs zur Wissenskonstruktion haben. Wird aber bei der Entwicklung von Lernaufgaben die Autonomie der Lernenden durch konkrete und kleinschrittige Aufgabenstellungen zu stark eingeschränkt, so kommt dieser Organisation in kleine Lernschritte ein behavioristischer Charakter zu. Auch die Idee der positiven Rückmeldungen bei AC für korrekt gebaute Moleküle stammt aus dem Behaviorismus, da sie die Funktion eines positiven Verstärkers haben. AC hat durchaus das Potenzial eines rein konstruktivistischen Lernsystems, je nach Anwendung können aber andere Lernmodelle einbezogen werden.

2.6 Das Lernsystem Kugel-Stab-Modell

Beim Vergleichsmedium zu AC handelt es sich um einen physischen Molekülbaukasten. Dieses Lernsystem wird in vorliegender Studie *Kugel-Stab-Modell* (KSM) genannt. Ein solcher Molekülbaukasten beinhaltet verschiedenfarbige Kugeln, welche die verschiedenen Atome repräsentieren. Jede Atom-Kugel ist mit derjenigen Anzahl Löcher versehen, wie das entsprechende Atom Bindungen eingehen kann. Die Bindungen sind durch Plastikstäbchen repräsentiert (Abbildung 2.11). Mit diesen Stäbchen und Kugeln kann man nun beliebige Moleküle zusammensetzen. Einige Beispielmoleküle finden sich in Abbildung 2.12.



Abb. 2.11: Der Molekülbaukasten: Kugel-Stab-Modell (KSM).

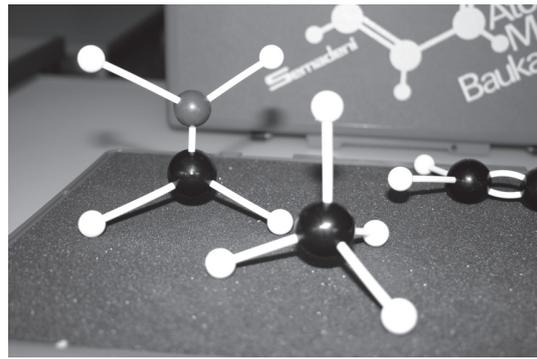


Abb. 2.12: Molekülmodelle nach dem Prinzip des KSM.

2.7 Unterschiede und Gemeinsamkeiten von Augmented Chemistry und dem Kugel-Stab-Modell

Im Folgenden wird kurz auf die Unterschiede und Gemeinsamkeiten von AC und dem KSM eingegangen.

Was die Veranschaulichung der Molekülmodelle betrifft, unterscheiden sich die beiden Systeme nicht. Beide sind dreidimensional repräsentiert. Bei AC können aber zusätzlich zur Darstellungsform mit Kugeln auch die chemischen Bezeichnungen (C, O, H, etc.) angezeigt werden. Auch unterliegen den beiden Systemen dieselben didaktischen Konzepte. Abgesehen davon dass AC einige Zusatzfunktionen bietet, liegt der grundsätzlichsste Unterschied der beiden Systeme lediglich in der Darstellungsform (virtuell bei AC, real beim KSM). AC bietet im Vergleich zum KSM zusätzlich folgende Funktionen: auditiv präsentierte Lerntexte zu den einzelnen Molekülen, Wechseln zwischen den Darstellungsformen *Ball-Stick* und *Wireframe* (siehe Abbildungen 2.13 und 2.14.), Anzeigen der Dipolmomente und die *Library*-Funktion, mit der man eine Reihe vordefinierter Moleküle durchsehen und von allen Seiten betrachten kann.



Abb. 2.13: Moleküldarstellung: *Ball-Stick*.



Abb. 2.14: Moleküldarstellung: *Wireframe*.

Ein weiterer Unterschied bietet sich in der Darstellungsform der spezifischen Informationen zu den gebauten Molekülen. Bei AC erhalten die untersuchten Personen Informationen auditiv, beim KSM liegen sie visuell als Lernkarten vor.

Momentan bietet das KSM insbesondere noch jenen Vorteil gegenüber AC, dass gleichzeitig verschiedene Moleküle betrachten und miteinander verglichen werden können.

2.8 Abschliessende Bemerkungen

Die vorliegende Evaluationsstudie erfolgt sowohl summativ zur Überprüfung des Endergebnisses von AC, als auch formativ mit dem Ziel Schwachstellen von AC zu finden und Verbesserungsmöglichkeiten herauszuarbeiten (vgl. Kap. 2.3.1). Auch wenn es sich bei AC um ein vollständig entwickeltes Programm handelt, liegt noch grosses Potenzial in Weiterentwicklungen und Verbesserungen der Massnahme.

Die Evaluation der Gebrauchstauglichkeit und Akzeptanz von AC stützt sich in vorliegender Arbeit auf den benutzerzentrierten Ansatz der ISO-Norm 9241-11 (siehe Kap. 2.3.3). Dazu werden die Usability-Kriterien Effektivität und Zufriedenheit genauer untersucht.

Aufgrund der in Kapitel 2.3.6 dargestellten Einschränkungen von Evaluationsstudien hat die vorliegende Arbeit nicht zum Ziel verallgemeinernde Aussagen über computergestützte interaktive Lernsysteme, die ähnliche Technologien wie AC verwenden, zu treffen. Das Ziel liegt vielmehr darin, ein besseres Verständnis der entscheidenden Faktoren für AC als erfolgreiche Lernumgebung zu erlangen.

Da es sich bei vorliegender Untersuchung um eine Vergleichsstudie handelt, können Aussagen über AC immer nur im Vergleich zum herkömmlichen Lernsystem zur Visualisierung von Molekülstrukturen, dem KSM, gemacht werden.

Wie empirische Studien gezeigt haben, herrscht Uneinigkeit und eine widersprüchliche Befundlage, was die Effektivität computergestützter Lernsysteme betrifft. AC ist also nicht automatisch besser als andere Lernmittel nur weil es interaktiv ist. Es ist wichtig neue Lernsysteme bezüglich ihrer Wirkung genau zu evaluieren, um ihren Erfolg auch zu garantieren. Das Ziel vorliegender Arbeit liegt darin, das Potenzial von AC zu erforschen und allfällige Probleme aufzudecken, um das System weiter zu optimieren.

3. Fragestellungen

1. Hat die Art der Lernmethode (Augmented Chemistry oder Kugel-Stab-Modell) einen Einfluss auf die Lernwirksamkeit?
 - 1.1 Sind bezüglich der praktischen Problemlösefähigkeit mit dem Lernsystem Unterschiede zwischen Personen, die mit Augmented Chemistry (AC), und solchen, die mit dem Kugel-Stab-Modell (KSM) arbeiten, erkennbar?
 - 1.2 Sind bezüglich des Lernerfolgs Unterschiede zwischen Personen, die mit AC lernen, und solchen, die mit dem KSM lernen, erkennbar?
2. Hat die Art der Lernmethode einen Einfluss auf das subjektive Empfinden bei den Lernenden?
 - 2.1 Unterscheiden sich Personen, die mit AC arbeiten, von solchen, die mit dem KSM arbeiten, hinsichtlich ihrer empfundenen Beanspruchung bei der Aufgabenbearbeitung?
 - 2.2 Gibt es hinsichtlich der emotionalen Stimmung nach Beendigung der Aufgabenbearbeitung Unterschiede zwischen Personen, die mit AC, und solchen, die mit dem KSM lernen?
 - 2.3 Sind bezüglich der Zufriedenheit mit der Lernsituation Unterschiede zwischen Personen, die mit AC, und solchen, die mit dem KSM lernen, erkennbar?
 - 2.4 Welches ist das von den Lernenden präferierte Lernsystem?
3. Gibt es Zusammenhänge zwischen der Lernwirksamkeit (praktische Problemlösefähigkeit und Lernerfolg) und der Beanspruchung sowie der emotionalen Stimmung?
 - 3.1 Unterscheiden sich die Zusammenhänge zwischen der Lernwirksamkeit und den subjektiven Empfindungen der Lernenden in der AC-Lernsituation von jenen in der KSM-Lernsituation?
4. Wie gut ist die subjektiv empfundene Usability von Augmented Chemistry?
5. Worin liegt das Potenzial von Augmented Chemistry als erfolgreiche und wirksame Lernumgebung?

4. Methode

Im Folgenden wird die Datenbasis und das methodische Vorgehen konkreter beschrieben. Vorweg wird aber nochmals kurz auf das Lernsystem AC eingegangen und beschrieben, in welcher Form es in dieser Studie eingesetzt wurde. Nacheinander werden dann die Stichprobe, der Versuchsplan, die Operationalisierung der verwendeten Variablen, die Versuchsdurchführung und zum Schluss die wichtigsten Auswertungsschritte beschrieben.

4.1 Einschränkungen für das Augmented Chemistry-System

Nach einer ausführlichen Beschreibung in Kapitel 2.5 wird in diesem Kapitel nochmals kurz auf das Lernsystem AC eingegangen. Da aus Kostengründen nicht das in Kapitel 2.5 beschriebene Originalsystem eingesetzt werden konnte, wird an dieser Stelle das System, wie es für die vorliegende Studie zur Verfügung stand, genauer beschrieben.

Der grosse Unterschied zwischen Original- und eingesetztem System liegt lediglich im Bildschirm. Für die Studie konnten nicht die teuren Bildschirme aus Acrylglas eingesetzt werden. Stattdessen wurden konventionelle 17" Röhrenmonitore verwendet. Es konnte somit auch nicht mit der Rückprojektionstechnik via Beamer gearbeitet werden (siehe Kapitel 2.5.2, Abbildung 2.6). Die Kamera wurde vorne an die verwendeten Bildschirme montiert. Neben der deutlich kleineren Bildfläche beim eingesetzten im Vergleich zum Originalsystem stellen die Röhrenmonitore eine weitere Einschränkung dar. Flachbildschirme wären aufgrund der besseren Bildqualität vorzuziehen gewesen, standen aber leider nicht zur Verfügung. Abbildung 4.1 zeigt die Ausstattung einer AC-Arbeitsstation wie sie in der vorliegenden Studie eingesetzt wurde sowie die Anordnung der einzelnen Arbeitswerkzeuge: Würfel, Buch, Greifer, Karten und Plattform.

Die AC-Stationen wurden nahe den Fenstern platziert, da natürliches Licht für optimale Verhältnisse sorgt. Die in den Schulzimmern vorhandene künstliche Beleuchtung führt nämlich zu Bildschwankungen. Einzig starkes Halogenlicht kommt als mögliche Lichtquelle neben dem Tageslicht in Frage. Bei den Versuchen wurde ausschliesslich mit Tageslicht gearbeitet. Da AC auch akustische Signale und auditiv präsentierte Informationen liefert sind die AC-Stationen mit Tonanlagen ausgestattet.

Die AC-Software hat des Weiteren die Funktionalität sogenannte *log files* zu erstellen. *Log files* sind Protokolldateien, welche jegliches Benutzerverhalten aufzeichnen. Im Fall von AC liefern diese *log files* Informationen beispielsweise darüber, welche und wieviele Moleküle gebaut wurden, wieviel Zeit dazu benötigt wurde und wieviele Fehler gemacht wurden.

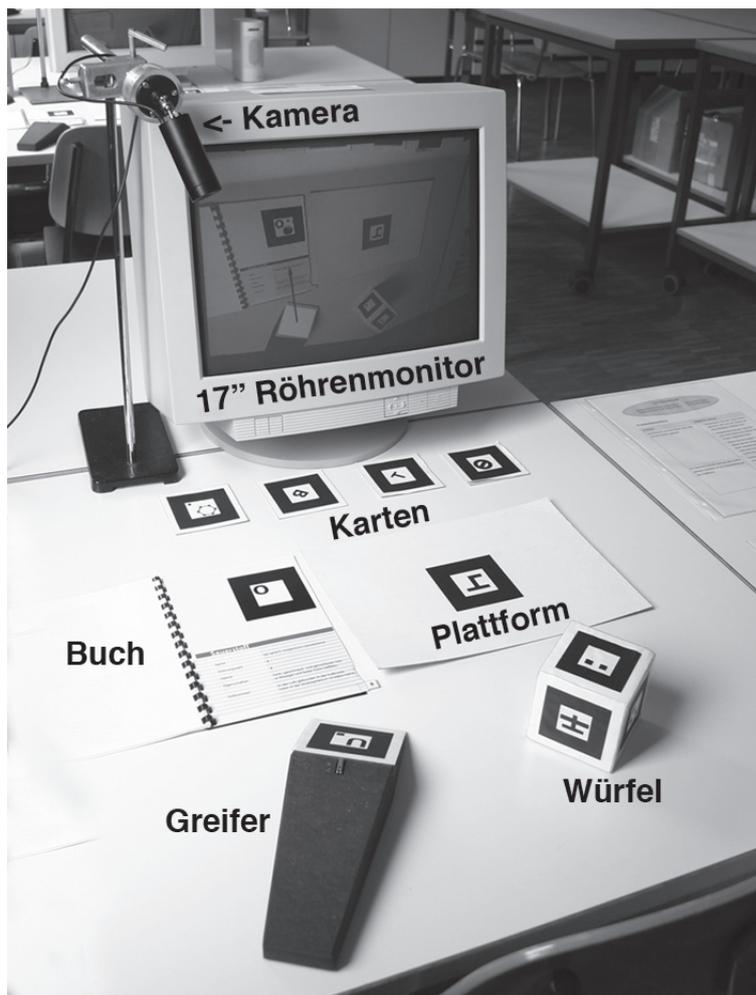


Abb. 4.1: Augmented Chemistry - Arbeitsplatz.

4.2 Stichprobe

An der Untersuchung nahmen 26 Lehrlinge der Berufsschule *aprentas* in Muttenz teil. *aprentas* ist ein Ausbildungszentrum für naturwissenschaftliche, technische und kaufmännische Berufe, insbesondere in der chemischen und pharmazeutischen Industrie. Weitere Informationen finden sich auf der Homepage www.aprentas.com. Bei den 26 Untersuchungspersonen (Upn) handelt es sich um zwei Biologielaborantenklassen im ersten Ausbildungsjahr mit einem Pensum an drei respektive vier Lektionen Unterricht in Chemie pro Woche. Die Stichprobe setzt sich aus 5 Frauen und 21 Männern zwischen 15 und 22 Jahren zusammen. Das durchschnittliche Lebensalter der Upn betrug 17.4 Jahre (SD = 1.8). Die Upn verbrachten zwischen einer und 16 Stunden am Computer pro Woche, was eine durchschnittliche Computernutzung von 6.9 Stunden pro Woche in der untersuchten Stichprobe ausmacht (SD = 4.1). Die Upn hatten zum Untersuchungszeitpunkt noch keine oder nur geringe Vorkenntnisse in Organischer Chemie und entsprachen somit der Zielgruppe des Lernsystems AC.

Die Untersuchung wurde angekündigt, indem die Lehrlinge vor Beginn der Untersuchung über Art und Umfang der Massnahme informiert wurden. Die Teilnahme an der Untersuchung war nicht freiwillig. *aprentas* hat sich bereit erklärt, AC als neuartiges Lernmedium auszuprobieren und versuchsweise in den Unterricht zu integrieren.

4.3 Versuchsplan

Die Evaluation, ob AC lernwirksam und zufriedenstellend eingesetzt werden kann, erfolgte mit einem Vergleichstestverfahren. Dabei ging es darum zu analysieren, ob AC vergleichbar oder sogar wirksamer ist als das herkömmliche, bei *aprentas* eingesetzte Lernmedium zur Visualisierung von Molekülstrukturen, das KSM. Die Art des Lernsystems - AC oder KSM - bildet folglich die unabhängige Variable, die auf Lernwirksamkeit und Akzeptanz als abhängige Variablen wirkt. Das Versuchsdesign ist in Abbildung 4.2 ersichtlich und wird im Folgenden erläutert.

Mögliche Störvariablen wie Vorwissen (Selbsteinschätzung), Motivation oder Einstellung zu computergestützten Lernmedien wurden in einem Vorfragebogen erfasst und zur Parallelisierung der 26 Upn in zwei möglichst gleichgrosse Lerngruppen herangezogen. Die eine Gruppe erarbeitete dabei den Lerninhalt mit AC, die andere Gruppe den identischen Lerninhalt mit dem KSM. Die Upn wurden pro Klasse den zwei Lerngruppen zugeordnet. So entstanden vorerst vier Gruppen (2 AC-Gruppen und 2 KSM-Gruppen), welche anschliessend zu einer AC- und einer KSM-Gruppe zusammengefasst wurden. Nach einer Lernphase absolvierten die Lernenden einen Kurztest und beurteilten danach die Lernsysteme anhand einiger Fragebögen. Damit alle Upn das neue Lernsystem AC kennenlernen durften, wurde eine Woche nach diesem Versuch ein zweiter Versuch mit neuem Lerninhalt durchgeführt. In der ersten Woche bearbeitete also die Gruppe 1 den Themenbereich A (Prinzipien des Aufbaus organischer Moleküle und Isomerie) mit AC, in der zweiten Woche den Themenbereich B (verschiedene Vertreter funktioneller Gruppen) mit dem KSM. Die Gruppe 2 verfuhr genau umgekehrt und bearbeitete in der ersten Woche Themenbereich A mit dem KSM, in der zweiten Woche Themenbereich B mit AC. Diese Versuchsanordnung ermöglichte eine zusätzliche Befragung der Upn nach der Präferenz für ein Lernsystem. Die Untersuchung verlief über einen Zeitraum von zwei Wochen.

In Kapitel 4.3.1 wird dieses spezielle Design noch genauer erläutert.

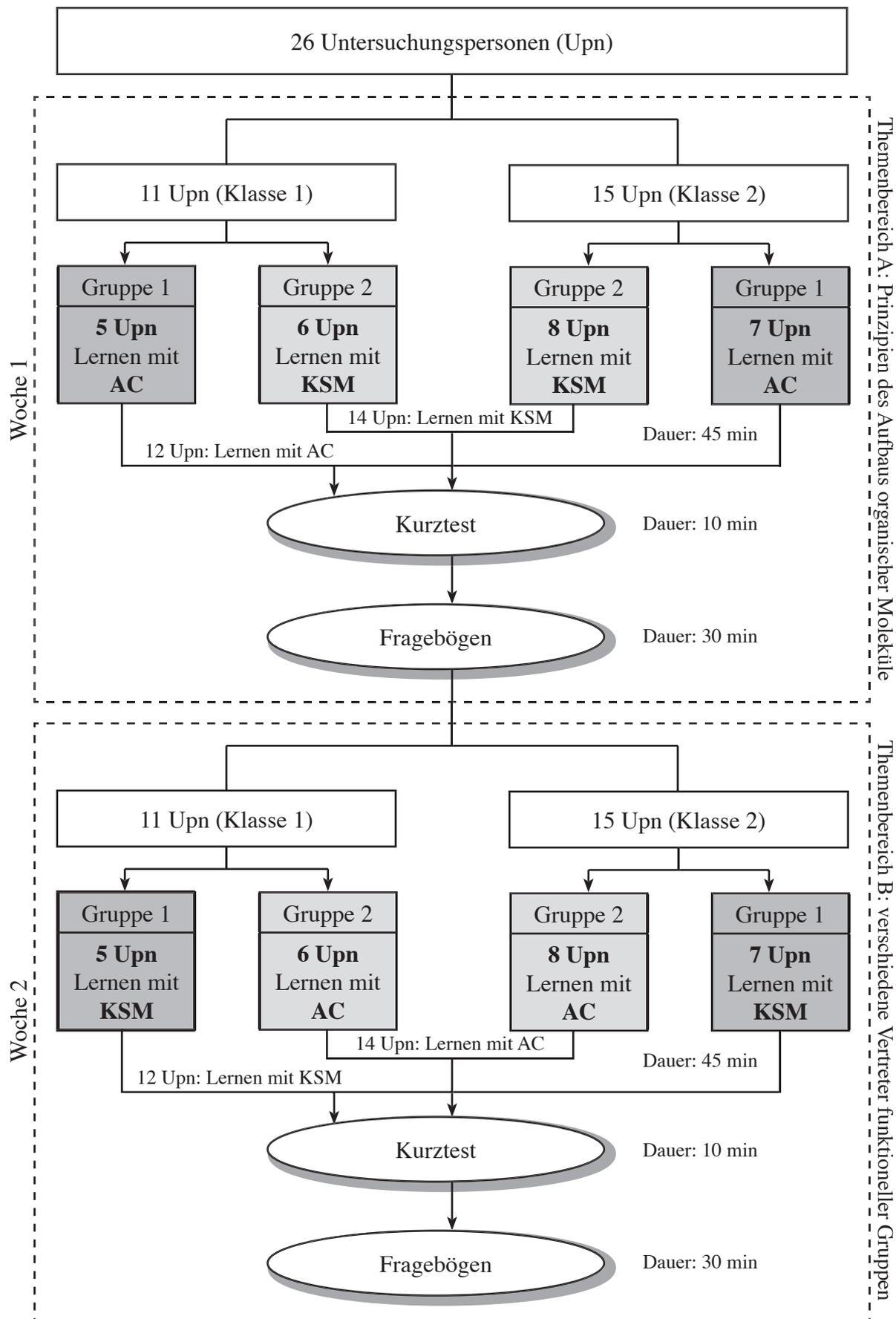


Abb. 4.2: Das Versuchsdesign.

4.3.1 AB/BA-Design

In Tabelle 4.1 wird das Versuchsdesign aus Abbildung 4.2 vereinfacht dargestellt. Dieser Versuchsaufbau entspricht dem einfachsten Design eines Crossover-Versuchs, einem 2×2 Design oder *two treatment two period crossover design* (vgl. Toutenberg, 1994; Jones & Kenward, 2003). Im Folgenden wird aber die häufig synonym verwendete Bezeichnung *AB/BA-Design* verwendet. Bei einem AB/BA-Design geht es grundsätzlich darum, dass zwei Behandlungen zu zwei Zeitpunkten getestet werden und zwar anhand von zwei Gruppen, welche die Behandlungen in unterschiedlicher Reihenfolge erhalten. Die beiden Behandlungen entsprechen in vorliegender Studie also den beiden Lernmethoden AC und KSM.

Tab. 4.1: Behandlungsabfolge in einem AB/BA-Design.

	Periode 1 (Thema A)	Periode 2 (Thema B)
Gruppe 1	Lernen mit AC	Lernen mit KSM
Gruppe 2	Lernen mit KSM	Lernen mit AC

Anmerkung: Periode 1 entspricht dem ersten Versuch in Woche 1, Periode 2 dem zweiten Versuch in Woche 2. Die beiden Versuche unterscheiden sich im Lerninhalt (Thema A und Thema B)

Die Gruppe 1 lernte folglich in Periode 1 mit AC und in Periode 2 mit dem KSM, während die Gruppe 2 genau umgekehrt zuerst mit dem KSM und dann mit AC lernte. Alle Upn testeten also beide Lernsysteme, wodurch der Vergleich der beiden Lernmethoden sozusagen “innherhalb” der Upn erfolgte. Jede Up war somit gleichzeitig Test- und Kontrollperson. So kann die Anzahl der benötigten Teilnehmer gegenüber anderen Studientypen mindestens halbiert werden. Das AB/BA-Design nutzt den Sachverhalt, dass die Variabilität von wiederholten Messungen an derselben Person (within-subject) geringer ist als die Variabilität von Messungen zwischen verschiedenen Personen (between-subject). Durch das AB/BA-Design können Effekte der Reihenfolge, mit der mit den beiden Lernsystemen gearbeitet wird, ausgeschlossen werden. Neben diesen Vorteilen, hat ein AB/BA-Design auch einen entscheidenden Nachteil. Es kann sein, dass der Effekt einer Lernmethode in Periode 1 bis in die nachfolgende Periode hineinwirkt. Wenn das Ergebnis der zweiten Behandlungsperiode durch das der ersten beeinflusst wird, spricht man von einem *Carry-Over-Effekt*, auch Nacheffekt genannt. Die möglichen Nachwirkungen in Periode 2 aus der vorgängigen Periode 1 können sich je nach Vorbehandlung (Lernen mit AC oder Lernen mit dem KSM) unterschiedlich auswirken. In einem AB/BA-Design muss aber vorausgesetzt werden, dass der Effekt einer Behandlung für beide Perioden derselbe ist. Liegt ein Carry-Over-Effekt vor, so erschwert oder gar verunmöglicht dies die Schätzung von direkten Behandlungseffekten (Toutenberg, 1994). Generell wird bei Carry-Over-Effekten geraten,

zur Aufdeckung von Behandlungsunterschieden nur Daten der ersten Periode zu verwenden, entsprechend einem between-subjects-Design (Lehmacher, 1987; Reed, 2003).

Die Messwerte der abhängigen Variablen für die insgesamt vier möglichen Kombinationen von Gruppe und Periode (siehe Tabelle 4.1) hängen im Allgemeinen von den unmittelbaren Effekten der beiden Lernmethoden (Behandlungseffekt), von generellen zeitlichen Veränderungen zwischen Periode 1 und Periode 2 (Periodeneffekt) und möglicherweise von Nachwirkungen in Periode 2 aus der Vorbehandlung in Periode 1 (Carry-Over-Effekt) ab. Der Carry-Over-Effekt hat nur einen Einfluss auf die abhängigen Variablen der zweiten Periode.

Ziel der Auswertung eines AB/BA-Versuchs ist es, das Vorhandensein dieser drei Effekte zu testen und ihre Grösse zu schätzen. Da in vorliegender Studie Aussagen über die Wirkung der Lernmethoden gemacht werden sollen, steht die Überprüfung der Daten auf einen Behandlungseffekt im Vordergrund. Dafür müssen aber Carry-Over-Effekte ausgeschlossen werden können. Der Carry-Over-Effekt stellt also das grösste Problem dar, denn davon hängt die ganze Interpretierbarkeit der Daten ab. Die statistischen Verfahren zur Messung der genannten Effekte werden in Kapitel 4.6 genauer erläutert.

4.4 Untersuchungsvariablen und ihre Operationalisierung

In der vorliegenden Untersuchung wurden sowohl objektive Leistungsdaten, als auch subjektive Beurteilungen erhoben. Zusammengefasst kamen folgende Methoden der Datenerhebung zum Einsatz: Aufgabenblätter und Tests zur Überprüfung der Lernwirksamkeit und Befragung in Form von Fragebogen. In den folgenden Unterkapiteln wird die Operationalisierung und Messmethode der verwendeten Variablen genauer erläutert.

4.4.1 Lernwirksamkeit

Die Messung der Lernwirksamkeit erfolgte über die zwei allgemeinen Kategorien Können und Wissen (Seel, 2000). Können wurde über die Variable *praktische Problemlösefähigkeit* erfasst, welche über Lernaufgaben operationalisiert wurde. Die Kategorie Wissen wurde der Variable *Lernerfolg* zugeschrieben, mit welcher der Aufbau deklarativen und prozeduralen Wissens überprüft werden sollte. Operationalisiert wurde der Lernerfolg über einen Kurztest.

Im Folgenden werden die Messmethoden der genannten Variablen genauer erläutert.

Die praktische Problemlösefähigkeit (Lernaufgaben)

Die praktische Problemlösefähigkeit wurde über sogenannte Lernaufgaben gemessen. Anhand dieser Lernaufgaben mussten die Upn den Lerninhalt mit Hilfe der beiden Lernsysteme AC und KSM erarbeiten. Die Aufgaben wurden schriftlich und selbständig bearbeitet. Zusammengestellt wurden die Aufgaben von einer Lehrperson bei aprentas. Im ersten Versuchsdurchlauf (Periode 1) bearbeiteten die Upn Lernaufgaben, mit denen sie die Prinzipien des Aufbaus organischer Moleküle und der Isomerie erlernen mussten (Themenbereich A). Im zweiten Versuchsdurchlauf (Periode 2) lernten die Upn mit den Lernaufgaben einige Vertreter verschiedener Stoffgruppen kennen (Themenbereich B). Die beiden Aufgabenblätter bestanden jeweils aus acht bis neun Aufgaben, die sich inhaltlich sowohl auf Molekülbau und -struktur, als auch auf Eigenschaften der gebauten Moleküle bezogen. Die einzelnen Aufgaben gliederten sich in Teilaufgaben, so wie es in Abbildung 4.3 dargestellt ist.

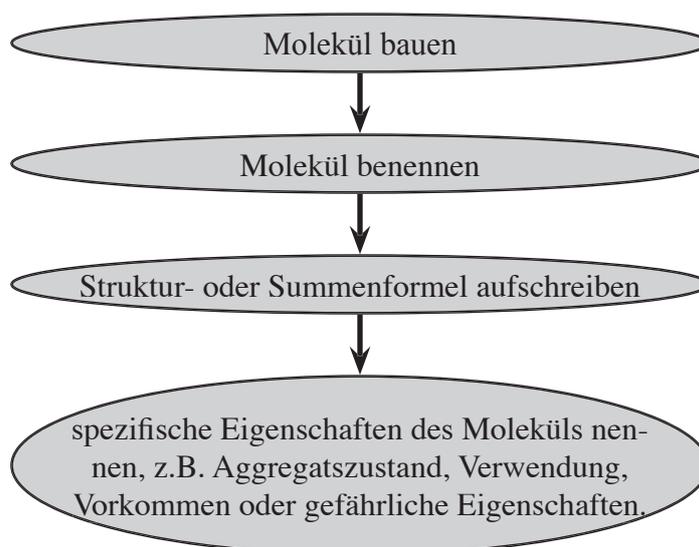


Abb. 4.3: Musterverlauf einer Lernaufgabe in den Aufgabenblättern.

Die beiden Versionen der Aufgabenblätter sind im Anhang A vollständig abgebildet. Während für den Themenbereich A einfache Kohlenwasserstoff-Moleküle wie Methan, Ethan oder Propan gebaut werden mussten, beinhalteten die Aufgaben zum Themenbereich B die Konstruktion komplexerer Moleküle wie Methanol, Aceton oder Essigsäure. Zur Bearbeitung der Lernaufgaben hatten die Upn jeweils 45 Minuten Zeit. Die erreichte Punktzahl in den Lernaufgaben dient als ein Mass dafür, wie gut mit AC oder dem KSM die entsprechenden Lerninhalte erarbeitet werden konnten. Erhoben wurde also die Fähigkeit der untersuchten Personen, in den entsprechenden Situationen selbständig Probleme und Aufgaben mit Hilfe der Lernsysteme zu

lösen bzw. zu bearbeiten. Deshalb wird die Variable zur Erfassung der Qualität der bearbeiteten Lernaufgaben unter dem Begriff *praktische Problemlösefähigkeit* zusammengefasst. Da die einzelnen Aufgaben in der Regel nicht ohne Hilfe der Lernsysteme gelöst werden konnten, sollte diese Problemlösefähigkeit wiederum aussagen, wie gut sich die Upn im Umgang mit dem jeweiligen Lernsystem zurecht fanden. Mit den Lernaufgaben wurde also auch eine praktische Kompetenz erfasst. Das durch die Lernaufgaben hervorgerufenen Lernergebnis wurde deshalb der Kategorie *Können* zugeordnet (nach Seel, 2000). Die Lernaufgaben wurden so bewertet, dass jede richtig gelöste Teilaufgabe einen Punkt ergab. Die Ergebnisse in den Lernaufgaben wurden schliesslich auf einer Skala von 0 (kein Erfolg) bis 100 (maximaler Erfolg) abgebildet.

Der Lernerfolg (Kurztest)

Der Lernerfolg wurde mit einem Kurztest gemessen. Dieser Kurztest schloss sich jeweils unmittelbar an eine Lerneinheit mit AC bzw. dem KSM an und prüfte den Inhalt der vorangegangenen Lerneinheit. Die beiden Kurztests zum Themenbereich A und B sind vollständig im Anhang B abgebildet. Die Tests beschränkten sich auf die Erfassung der Behaltensleistung durch Reproduktion oder Wiedererkennen von Wissensinhalten aus den Lernaufgaben. Es ging vordergründig darum, wieviel die Upn nach der Lerneinheit inhaltlich noch wussten. Die einzelnen Testaufgaben setzen sich aus folgende Aufgabentypen zusammen:

- Einen Stoff anhand der Summen- oder Strukturformel benennen
- Die Strukturformel eines Moleküls zeichnen
- Summenformel eines Moleküls angeben
- Aggregatzustand eines Moleküls angeben
- Spezifische Eigenschaften eines Moleküls aufzählen

Der Kurztest prüfte, wieviel Wissen die Upn in den Lerneinheiten erworben hatten, und zwar wurde damit sowohl deklaratives als auch prozedurales Wissen (beim Zeichnen von Strukturformeln) abgefragt. Dieses erworbene Wissen als Lernergebnis wird unter dem Begriff *Lernerfolg* zusammengefasst.

Zum Lösen der Testaufgaben hatten die Upn jeweils 10 Minuten Zeit. Die Testaufgaben wurden so bewertet, dass jede richtig gelöste Aufgabe einen Punkt ergab. Die Ergebnisse der Kurztests wurden schliesslich auf einer Skala von 0 (kein Erfolg) bis 100 (maximaler Erfolg) abgebildet.

4.4.2 Beanspruchung

Zur Messung des Beanspruchungszustands der Lernenden bei der Aufgabenbearbeitung wurde der NASA Task Load Index von Hart und Staveland (1988) in deutscher Version (vgl. Unema et al., 1988) eingesetzt. Der NASA Task Load Index (NASA-TLX) ist eine mehrdimensionale subjektive Beanspruchungsmessmethode und erfasst die Beanspruchung anhand folgender sechs Subskalen:

- Geistige Beanspruchung
- Körperliche Beanspruchung
- Zeitliche Beanspruchung
- Anstrengung
- Aufgabenbewältigung
- Frustration

Die Skalen sind 20-stufig und die Ratings werden in einen Wertebereich von 0 ('gering Beanspruchung') bis 100 ('hohe Beanspruchung') überführt. Die Bewertung der Beanspruchung ist in zwei Phasen unterteilt. Zunächst mussten die Upn ihre Lernsituation hinsichtlich der erlebten Beanspruchung in den sechs Bereichen bewerten. Der genaue Wortlaut der Fragen zu jeder Dimension ist dem Fragebogen zu entnehmen, welcher in Anhang C abgedruckt ist. Danach wurden die Subskalen von den Upn gewichtet. Die Gewichtung erfolgt im NASA-TLX über einen Paarvergleich der sechs Skalen, bei dem die Up bei jedem der zu vergleichenden Paare angeben muss, welcher der beiden Skalen sie grössere Bedeutung für die Beanspruchungsbewertung der gegebenen Situation beimisst. Es wurde also in einem ersten Schritt das Beanspruchungsprofil der Lernsituation bestimmt, welches anschliessend unter dem Gesichtspunkt der Bedeutung der einzelnen Subskalen für die Gesamtbeanspruchung gewichtet wurde. Zum Schluss wurden die einzelnen Ratings mit den Gewichtungen zu einem gewichteten Gesamtwert verrechnet, der die Beanspruchung in Prozent angibt. Dieser gewichtete Gesamtwert der Beanspruchung entspricht dem *Overall Weighted Workload Score* der englischen Version.

Im englischen Sprachraum hat sich der NASA-TLX im Vergleich zu anderen subjektiven Beanspruchungsskalen bisher am besten bewährt. Was die Reliabilität und Validität der NASA-TLX-Skalen angeht, sind für den deutschen Sprachraum kaum Informationen vorhanden. Pfendler (1990) findet aber Hinweise auf eine gute Reliabilität des gewichteten Gesamtwertes und der Subskalen. Hingegen bestehen Hinweise auf eine niedrigere Validität der Messmethode im Vergleich zu anderen subjektiven Verfahren.

4.4.3 Emotionale Stimmung

Für eine emotionale Beurteilung der Lernsysteme wurde der Self Assessment Manikin (SAM) von Lang (1985) eingesetzt (Abbildung 4.4). Mit dem SAM wurde die emotionale Reaktion der Upn auf die Lernsysteme AC und KSM ermittelt.

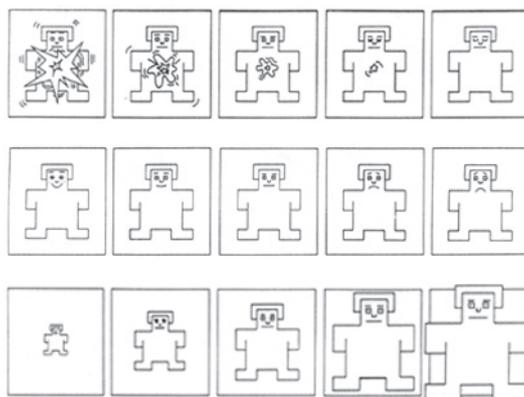


Abb. 4.4: Bildvorlage im Self Assessment Manikin (SAM) zur Beurteilung der emotionalen Dimensionen *Aktiviertheit* (obere Bildpalette), *Valenz* (mittlere Bildpalette) und *Dominanz* (untere Bildpalette).

Beim SAM handelt es sich um ein nicht-sprachliches Verfahren der Emotionsmessung, das die Ausprägungen auf den drei affektiven Dimensionen *Aktiviertheit*, *Valenz* und *Dominanz* erfasst. Mit diesem dreidimensionalen Modell der Emotionen stützt sich Lang (1985) auf bisherige Befunde, bei welchen immer wieder eine 3 Faktorenlösung herauskam (z.B. Osgood, Suci & Tannenbaum, 1957; Mehrabian & Russell, 1974). Während die beiden Dimensionen Aktiviertheit und Valenz heute praktisch unbestritten sind, wird die Frage nach einer dritten Dimension sehr kontrovers diskutiert. Trotz dem umstrittenen Vorliegen einer dritten Dimension wird Dominanz als solches in die Untersuchung mit einbezogen. Diese Dimension ist gerade im Zusammenhang der Bewertung von Softwareprodukten ein entscheidendes Kriterium. Gehört es doch zu einem der Grundsätze einer benutzerfreundlichen Mensch-Maschine-Interaktion, dass der Benutzer die Software und nicht Software ihn kontrolliert (vgl. Hegner, 2003, S. 13). Genau dieser Aspekt wird versucht mit der emotionalen Dimension Dominanz zu erfassen.

Beim SAM stufen die Befragten anhand von Figuren, deren "Innenleben", Gesichtsausdruck oder Grösse variiert, ihren Gefühlsausdruck mit einem Kreuz auf oder zwischen den Figuren ein. Dadurch entsteht eine 9-stufige Skala. Die Aktiviertheit wird durch eine Reihe von Figuren gemessen, welche von einer aufgeregten mit weit geöffneten Augen ('emotional erregter Zustand') bis hin zu einer sehr ruhigen, fast schlafenden Figur übergeht ('ruhiger Zustand').

Die Dimension Valenz wird dadurch repräsentiert, dass die Figuren ihren Gesichtsausdruck von einem Lächeln ('angenehmes Gefühl') bis zu einem traurigen, missgestimmten Ausdruck ('unangenehmes Gefühl') verändern. Die Dimension Dominanz erfasst, wie sehr die Upn das Gefühl haben, die Situation unter Kontrolle haben und wird durch die Zunahme der Grösse der Figuren repräsentiert (von 'submissiv' bis 'dominant'). Die Dimensionen wurden in vorliegender Studie so bewertet, dass entsprechend hohe Ausprägungen ein hohes Mass an Aktiviertheit, Valenz oder Kontrollempfinden ausdrücken. Auf der 9-stufigen Skala wurde die Aktiviertheit somit von 1 ('ruhig') bis 9 ('erregt'), die Valenz von 1 ('unangenehmes Gefühl') bis 9 ('angenehmes Gefühl') und die Dominanz von 1 ('submissiv') bis 9 ('dominant') bewertet.

Der SAM kommt im Vergleich zu den benötigten 18 Einzelbewertungen beim Semantischen Differential von Mehrabian und Russell (1974) mit nur drei Einzelbewertungen aus. Der SAM ist also ein sehr ökonomisches Instrument. Darüber hinaus ermöglicht dieses Verfahren die Messung von Emotionen unabhängig von eventuell missverständlichen verbalen Anker. Der SAM erlaubt zudem eine Erfassung der unmittelbaren emotionalen Reaktion auf die Lernmethoden, da der Fragebogen jeweils unmittelbar nach einer Lernsituation bearbeitet wurde. Die schnelle Beantwortung des SAM sollte die Erfassung der momentanen Gefühlslage ermöglichen, bevor dieser Gefühlszustand wieder verflogen ist. Der Fragebogen ist im Anhang D vollständig abgebildet.

4.4.4 Zufriedenheit mit der Lernsituation

Die Messung der Zufriedenheit der Upn mit den beiden Lernsystemen erfolgte über subjektive Daten, welche für die beiden untersuchten Lernsysteme miteinander verglichen werden konnten. Die Zufriedenheit der Upn mit den Lernsystemen wurde durch einen eigens konstruierten Fragebogen erfasst (Anhang E). In diesem Fragebogen mussten die Lernenden einschätzen, wie sie das Arbeiten mit AC oder dem KSM empfunden hatten. Der Fragebogen wurde also von den beiden Lerngruppen gleichermassen ausgefüllt. Somit konnte gemessen werden, wie sich die beiden Lernsysteme in ihrer Beurteilung unterschieden. Der Fragebogen enthält zum einen ein Semantisches Differential, anhand dessen die Upn die Lernsysteme durch 12 bipolare verbale Anker mittels einer 5-stufigen Ratingskala beschreiben mussten. Positive Pole der Beurteilungsskalen des Semantischen Differentials sind beispielsweise "interessant", "klärend" oder "effizient". Das Semantische Differential erbringt wesentliche Hinweise auf die generelle affektive Wirkung eines interaktiven Systems auf den Benutzer. Neben diesem Semantischen Differential beinhaltet der Fragebogen auch 12 Item-Aussagen zu bestimmten Aspekten der Lernsituation, bei welchen die Upn den Grad ihrer Zustimmung auf einer 5-stufigen Likertskala angeben mussten. Die Antworten in den 12 Item-Aussagen sollten Informationen über die Zufriedenheit der Upn unter bestimmten Gesichtspunkten der Lernsituation liefern. Beispielitems

sind: “Das Arbeiten mit Augmented Chemistry (oder dem Kugel-Stab-Modell) hat mir Spass gemacht” oder “Augmented Chemistry (oder das Kugel-Stab-Modell) hat mir geholfen, den Lerninhalt besser zu verstehen”.

Des Weiteren wurden offene Fragen gestellt, bei welchen die Lernenden positive und negative Aspekte der Lernsysteme selbst formulieren konnten. Bei AC durften zudem Verbesserungsvorschläge gemacht werden. Zum Schluss wurden die Upn gebeten, eine Gesamtbewertung der jeweiligen Lernsituation auf einer Schulnotenskala von 1 (“ungenügend”) bis 6 (“sehr gut”) anzugeben. Diese Bewertung wird als Mass für die allgemeine Zufriedenheit mit den Lernsystemen betrachtet, die sich aus den gesammelten Erfahrungen mit den verschiedenen Aspekten der beiden Lernsituationen ergeben hat.

4.4.5 Präferenzbeurteilung

Dank der Tatsache, dass in vorliegender Studie alle Upn beide Lernsysteme getestet hatten, konnte eine Befragung nach der Präferenz eines Lernsystems durchgeführt werden. Das von den Lernenden präferierte Lernsystem wurde mittels eines Fragebogens (siehe Anhang F) nach der zweiten Versuchsdurchführung erhoben. In diesem Fragebogen mussten die Upn anhand von neun ausgewählten Aspekten beurteilen welches Lernsystem ihnen beispielsweise mehr Spass bereitet hatte, leichter zu erlernen oder effizienter zum Lernen war.

4.4.6 Subjektive Usability von Augmented Chemistry

Die Benutzerzufriedenheit von AC als Softwareprodukt wurde mit dem Benutzerfragebogen SUMI (*Software Usability Measurement Inventory*) in deutscher Version von Kirakowski und Corbett (1993) erhoben (Anhang G). SUMI wurde von der *Human Factors Research Group* (HFRG) am University College Cork in Irland entwickelt und validiert (Porteous, Kirakowski & Corbett, 1993). Mit dem SUMI konnte die von den Benutzern subjektiv empfundene Usability von AC gemessen werden. SUMI ist ein standardisiertes Verfahren zur subjektiven Bewertung von Software durch Benutzer. Der Fragebogen umfasst 50 Item-Aussagen, die auf einer 3-stufigen Skala beurteilt werden müssen (einverstanden/weiss nicht/nicht einverstanden). Beispielitems sind: “Die Software hängt sich gelegentlich unerwartet auf” oder “Ich habe das Gefühl, die Software zu beherrschen, wenn ich sie benutze”. Die 50 Items können fünf Subskalen zugeordnet werden, welche mit ihrer genaueren Bedeutung der Tabelle 4.2 zu entnehmen sind.

Tab. 4.2: Die fünf Subskalen von SUMI und ihre Bedeutung.

Name (engl.)	Name (dt.)	Beschreibung
Affect	Emotionale Bewertung	Die generelle emotionale Reaktion der Benutzer auf die Software. Eine Bewertung wie angenehm, erfreulich und stimulierend der Umgang mit dem Produkt empfunden wird.
Control	Steuerbarkeit	Das Ausmass, in dem die Software konsistent auf Eingaben und Befehle der Benutzer reagiert und seine Handlungen vorhersehbar sind. Die Benutzer sollten die Software kontrollieren und nicht die Software sie.
Efficiency	Effizienz	Das Ausmass an Unterstützung, das die Benutzer bei ihrer Arbeit durch die Software erfahren und inwiefern eine effektive, schnelle und ökonomische Aufgabendurchführung möglich ist.
Helpfulness	Selbstbeschreibungsfähigkeit	Das empfundene Ausmass, mit dem die Software hilfreiche Anweisungen oder Hinweise gibt, insbesondere auch bei Problemen mit der Bedienung.
Lernability	Erlernbarkeit	Der benötigte Aufwand, um den Umgang mit der Software zu erlernen und das Gefühl zu bekommen, die Software zu beherrschen.

Anmerkung: die deutschen Übersetzungen der Skalennamen (siehe 2. Spalte) sind frei nach der Autorin vorgenommen.

Neben den fünf Subskalen existiert eine Globalskala, welche sich auf die generelle Empfindung von Software-Usability der Benutzer bezieht. Die Reliabilität (Cronbachs Alpha) der Subskalen wird auf einen Wert zwischen $\alpha = 0.71$ und $\alpha = 0.85$ und jene der Globalskala auf $\alpha = 0.92$ geschätzt. Das Verfahren ist zudem dafür geeignet Hinweise auf mögliche Normabweichungen zu erhalten. SUMI basiert nämlich auf einer umfassend standardisierten Datenbasis, welche anhand einer Vielzahl von Softwareprodukten angelegt wurde. So können Aussagen gemacht werden, in welchen Bereichen AC besser oder schlechter war als eine "Standardsoftware". SUMI dürfte momentan der psychometrisch am solidesten konstruierte Benutzerfragebogen zur globalen Bewertung von Dialogsystemen sein (Dzida et al., 2000; zit. nach Hegner, 2003, S. 59).

4.5 Versuchsdurchführung

Vor der eigentlichen Versuchsdurchführung wurden die Upn anhand eines Briefes ausführlich über die geplante Studie informiert. Zudem wurden die Upn vor den Versuchen vom Versuchsleiter in die Benutzung von AC eingeführt. Diese Einführung in AC wurde zuvor an einer Chemielaboranten-Klasse getestet und von ihr bewertet. Somit konnte überprüft werden, ob die

geplante Einführung ausführlich genug war, und ob sich die Benutzer danach imstande fühlten Aufgaben mit dem System zu lösen. Die beiden für die Studie untersuchten Klassen wurden gruppenweise in AC eingeführt. Nach einer ca. zehnmütigen Demonstration der einzelnen Bedienelemente und der Benutzung von AC durch den Versuchsleiter, hatten die Upn pro Gruppe während einer halben Stunde Zeit sich zu zweit ans AC-System zu setzen und den Umgang damit zu üben. Die Einführung in AC verlief nach einem standardisierten Ablauf, so dass für alle Upn gleiche Bedingungen vorausgesetzt werden konnten.

Aufgrund von Ausfällen des Schulbetriebs infolge eines Feiertags, konnte mit der ersten Versuchsdurchführung leider erst zwei Wochen nach der Einführung in AC begonnen werden. Für die beiden Versuchsdurchführungen wurden zweimal die Wochenlektionen in Chemie der beiden Klassen beansprucht. Aufgrund der Klassengrößen und der Tatsache, dass jeweils nur drei Personen gleichzeitig an den Lernsystemen arbeiten konnten, wurde ein Unterrichtsplan erstellt, der allen Upn während den benötigten Chemie-Lektionen eine Beschäftigung gab. Die Upn rotierten zwischen folgenden Stationen: Lerneinheit mit AC oder dem KSM, Kurztest und Fragebogen, Rahmenprogramm. Jede dieser drei Stationen fand in einem separaten Schulzimmer statt. Während die ersten sechs Upn (3 zu AC und 3 zum KSM) mit der Lerneinheit begannen, wurden die restlichen Upn mit dem Thema "Radioaktivität" beschäftigt. Dieses sogenannte Rahmenprogramm ist für die Studie inhaltlich nicht relevant aber notwendig, weil nicht alle gleichzeitig mit den Lernsystemen arbeiten konnten. Im Folgenden wird der Ablauf einer Lerneinheit zuerst für AC, dann für das KSM erläutert. Die beiden Lerneinheiten fanden parallel in separaten Räumen statt.

Die Lerneinheit mit AC

Die Versuchsanordnung ist in Abbildung 4.5 veranschaulicht. Für die Versuche standen drei AC-Arbeitsstationen zur Verfügung, welche gemäss Abbildung 4.5 im Raum angeordnet wurden. Der Versuchsleiter befand sich im selben Raum. Eine Lerneinheit wurde von jeder Up individuell absolviert. Dazu verteilten sich die jeweils drei Upn auf die vorhandenen drei AC-Stationen, danach wurde das Aufgabenblatt ausgehändigt, und die Upn hatten während 45 Minuten Zeit, das Aufgabenblatt mit Hilfe von AC zu lösen. Nach Abgabe des fertig gelösten Aufgabenblattes füllten die Upn den Self Assessment Manikin aus, worauf sie in die Pause entlassen wurden.

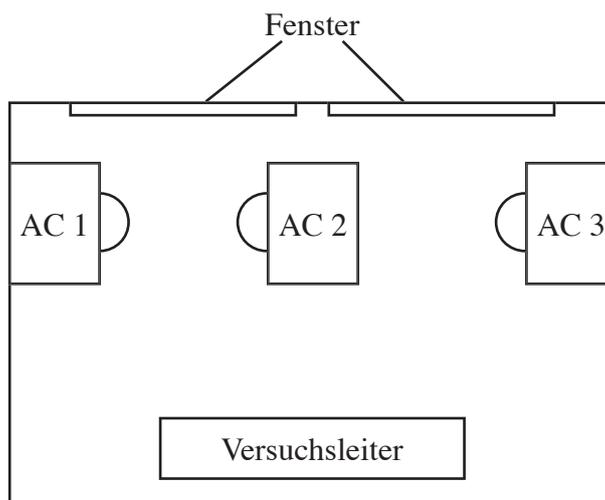


Abb. 4.5: Räumliche Anordnung der AC-Stationen im Versuchsraum.

Lerneinheit mit dem KSM

Für die Lerneinheit mit dem KSM wurden drei Molekülbaukästen eingesetzt, welche gemäss der Skizze in Abbildung 4.6 im Raum angeordnet waren. In demselben Raum befand sich zusätzlich eine Lehrperson. Der Ablauf der Lerneinheit wurde mit einer Videokamera aufgezeichnet.

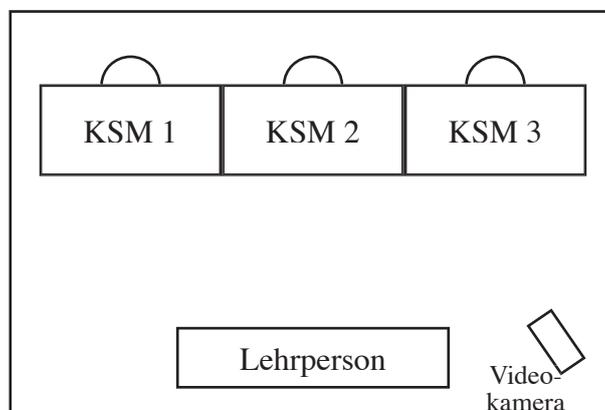


Abb. 4.6: Räumliche Anordnung der KSM-Baukästen im Versuchsraum.

Da die Lernenden weder Vorkenntnisse in Organischer Chemie noch Kenntnisse im Umgang mit dem KSM hatten, wurden zur Vereinfachung die für die Aufgabenstellung benötigten Elemente des Molekülbaukastens entsprechend beschriftet. Hierbei wurden die Atomkürzel (C, O, N) und die Art der Bindungen (Einfach- oder Doppelbindung), die eingegangen werden können, angegeben. Diese Hilfestellung war auch deshalb notwendig, weil bei AC die Bezeichnung der Atome und der Art der Bindungen durch das System gewährleistet wird.

Während sich die Bedingungen und der Lerninhalt einer KSM-Lerneinheit nicht von einer AC-Lerneinheit unterscheiden, so war das Lernen mit dem KSM hingegen an eine stärkere Interaktion mit einer Lehrperson gebunden. Die bei AC vom Computer übernommenen Funktionen wie das Benennen von Molekülen und die automatische Einspielung gesprochener Informationen zum Molekül, mussten bei der KSM-Lerneinheit auf andere Weise gewährleistet werden. So mussten die Upn beim KSM jeweils nach einem fertig gebauten Molekül die Lehrperson aufrufen, welche das Molekül nach seiner Richtigkeit beurteilte. Für ein richtig gebautes Molekül erhielten die Upn dann eine Lernkarte mit den benötigten Informationen für die Beantwortung der weiteren Fragen. Die Informationen auf diesen Lernkarten waren identisch mit den auditiv präsentierten Lerntexten bei AC. Dieses Vorgehen wiederholte sich für jedes gebaute Molekül.

In der Abbildung 4.7 wird anhand des bereits in Abbildung 4.3 veranschaulichten Musterverlaufs einer Aufgabenstellung für die beiden Lernsysteme AC und KSM vergleichend dargestellt, wie die Teilaufgaben gelöst werden konnten.

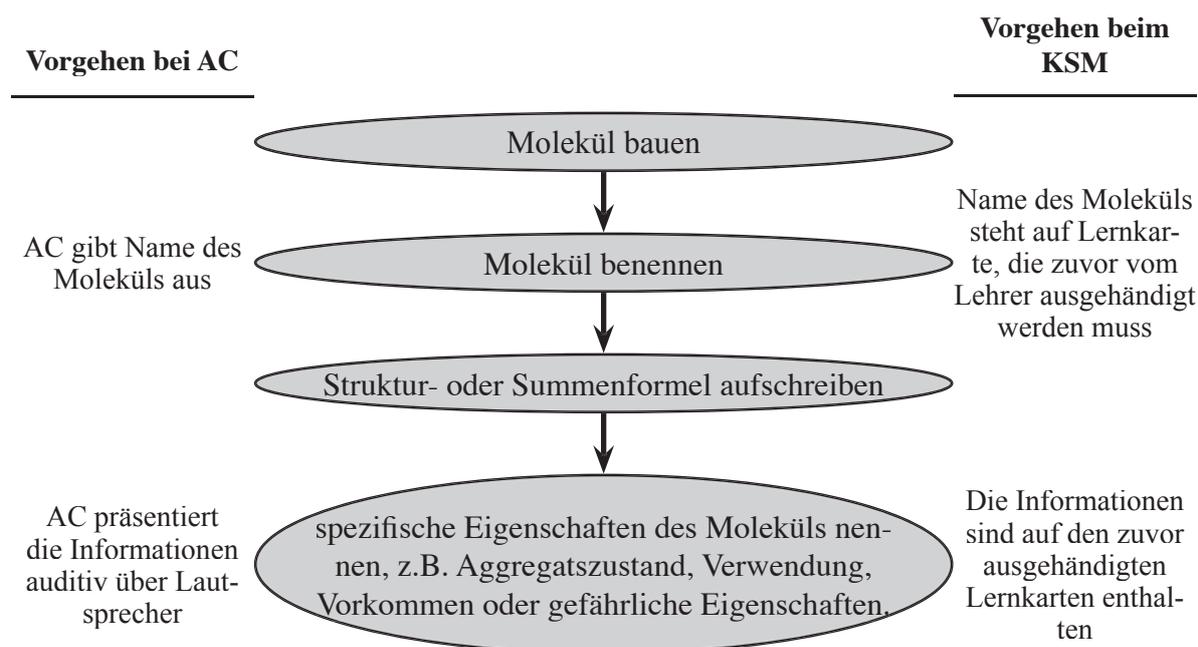


Abb. 4.7: Vergleichende Darstellung des Vorgehens bei der Aufgabenlösung mit AC und mit dem KSM.

Der im Anschluss an die Lerneinheit gestellte Kurztest fand unter Aufsicht und ohne Hilfsmittel statt. Nach dem Kurztest füllten die Upn zusätzlich die in den Kapiteln 4.4.2 bis 4.4.6 beschriebenen Fragebögen zur Bewertung des jeweiligen Lernsystems aus.

4.6 Auswertungsverfahren

Im Folgenden werden die wichtigsten Auswertungsmethoden erwähnt, mit deren Hilfe die Fragestellungen der vorliegenden Untersuchung beantwortet werden sollen.

An dieser Stelle sei nochmals auf den Grund und den Nutzen von zwei Versuchsdurchführungen in vorliegender Studie hingewiesen. Einerseits hatten somit alle Upn die Gelegenheit, das neue Lernsystem AC kennenzulernen. Andererseits konnte dadurch, dass alle Upn mit beiden Lernsystemen gearbeitet hatten, eine Befragung der Upn nach der Präferenz eines Lernsystems durchgeführt werden.

4.6.1 Auswertung des AB/BA-Designs

Die Fragen nach Unterschieden zwischen AC und dem KSM hinsichtlich der Lernwirksamkeit und der Beanspruchung, der emotionalen Stimmung und der Zufriedenheit bei den Lernenden werden über die Auswertung des AB/BA-Designs beantwortet. Die Auswertung des AB/BA-Designs erfolgt durch eine spezielle Form einer messwiederholten Varianzanalyse über F-Tests. Anhand der Varianzanalyse können die drei Effekte eines AB/BA-Designs (Behandlungseffekt, Periodeneffekt, Carry-Over-Effekt) getestet werden. In Tabelle 4.3 sind die drei unabhängigen Variablen aus dem AB/BA-Design abgebildet, welche jeweils die zweistufigen Faktoren des varianzanalytischen Designs ergeben.

Tab. 4.3: Faktoren der Varianzanalyse des AB/BA-Designs.

Faktor	Beschreibung
Gruppe (between-subject)	1: Abfolge AC/KSM; 2: Abfolge KSM/AC
Lernmethode (within-subject)	1: Lernmethode AC; 2: Lernmethode KSM
Periode (within-subject)	1: Versuch 1 (Woche 1); 2: Versuch 2 (Woche 2)

Es gibt zwei within-subjects Faktoren *Lernmethode* und *Periode* und einen between-subjects Faktor *Gruppe*. Die Periode entspricht dem Untersuchungszeitpunkt. Da AB/BA-Designs keine vollständige faktoriellen Pläne sind, kommen nicht alle, sondern nur vier mögliche Kombinationen der drei Faktoren vor. Alle Haupteffekte sind mit einer Wechselwirkung 1. Ordnung konfundiert. Dies bedeutet, dass jede Schätzung eines Haupteffekts gleichzeitig einen Interaktionseffekt zwischen den beiden übrigen Faktoren schätzt, also ein sogenanntes Alias hat. Die Schätzung eines Effekts kann folglich auf unterschiedliche Weise interpretiert werden. In Tabelle 4.4 werden die drei Haupteffekte mit ihren entsprechenden Interaktionseffekten abgebildet.

Tab. 4.4: Die Haupteffekte des varianzanalytischen Designs mit ihren entsprechenden Interaktionseffekten.

Haupteffekt	Interaktionseffekt (Alias)
Gruppe	Lernmethode*Periode
Lernmethode	Gruppe*Periode
Periode	Gruppe*Lernmethode

Falls einer der Haupteffekte signifikant ausfällt, kann man leider nicht feststellen, ob der Effekt selbst, sein Alias oder beide das signifikante Resultat hervorgebracht haben. Signifikante Effekte, welche mit dem Faktor Gruppe verknüpft sind, sind nicht zu erwarten, da die Upn durch Parallelisierung den beiden Gruppen zugeordnet wurden. So sollte ein signifikanter Haupteffekt der Gruppe über das Resultat der Interaktion zwischen Lernmethode und Periode interpretiert werden, und signifikante Haupteffekte der Behandlung bzw. der Periode sollte man nicht als das Resultat eines Interaktionseffekts zwischen Gruppe und Periode bzw. zwischen Gruppe und Lernmethode betrachten.

In Tabelle 4.5 ist abgebildet, über welche statistischen Effekte der Varianzanalyse die drei Effekte Behandlungs-, Perioden- und Carry-Over-Effekt interpretiert werden. Dabei sind diejenigen Effekte - ob Haupteffekt oder Interaktionseffekt - kursiv gedruckt, über welche die drei Effekte des AB/BA-Designs geschätzt werden sollen.

Tab. 4.5: Interpretationsgrundlage der drei Effekte (Behandlungs-, Perioden-, Carry-Over-Effekt) eines AB/BA-Designs aufgrund der Faktoren der Varianzanalyse.

Haupteffekt	Interaktionseffekt (Alias)	Effekt des AB/BA-Designs
Gruppe	<i>Lernmethode*Periode</i>	=> Carry-Over-Effekt
<i>Lernmethode</i>	Gruppe*Periode	=> Behandlungseffekt
<i>Periode</i>	Gruppe*Lernmethode	=> Periodeneffekt

Anmerkung: die Effekte des AB/BA-Designs werden über die kursiv gedruckten Effekte der Varianzanalyse interpretiert.

Gemäss Tabelle 4.5 wird der Carry-Over-Effekt über den Interaktionseffekt zwischen Lernmethode und Periode interpretiert. Ein Carry-Over-Effekt existiert also dann, wenn die Lernmethode je nach Periode einen anderen Effekt hat. Behandlungs- und Periodeneffekte werden über die Haupteffekte der Faktoren Lernmethode bzw. Periode interpretiert. Da die drei Effekte des AB/BA-Designs nur über die beiden Faktoren Lernmethode und Periode geschätzt werden können, erfolgen die Auswertungen über die zweifaktorielle Varianzanalyse.

Im Folgenden werden die Effekte des AB/BA-Designs erläutert und anhand von Interaktionsdiagrammen veranschaulicht. Dabei werden die Mittelwerte beider Gruppen gegen die Zeit aufgetragen und pro Lernmethode die Punkte miteinander verbunden. So wird die mittlere Tendenz der Daten im Hinblick auf Veränderungen in der Wirkung der jeweiligen Lernmethode deutlich.

Ein **Behandlungseffekt** liegt dann vor, wenn die Wirkung der beiden Lernmethoden AC und KSM verschieden ist. Sofern kein Carry-Over-Effekt vorliegt, tritt der Behandlungsunterschied in beiden Gruppen gleich stark auf. In Abbildung 4.8 ist ein reiner Behandlungseffekt dargestellt. Bei einem **Periodeneffekt** liegt das Mittel der beiden Behandlungswirkungen in der zweiten Periode auf einem anderen Niveau als das der ersten Periode. Solche Periodeneffekte sind bei den Variablen zur Lernwirksamkeit durchaus möglich, da sich die Aufgabenstellungen in den beiden Versuchen möglicherweise auch im Schwierigkeitsgrad unterscheiden. Sofern Carry-Over-Effekte ausgeschlossen werden können, sind diese Periodeneffekte aber unproblematisch und beeinflussen die statistische Analyse nicht. Abbildung 4.9 veranschaulicht einen Periodeneffekt in Verbindung mit einem Behandlungseffekt.

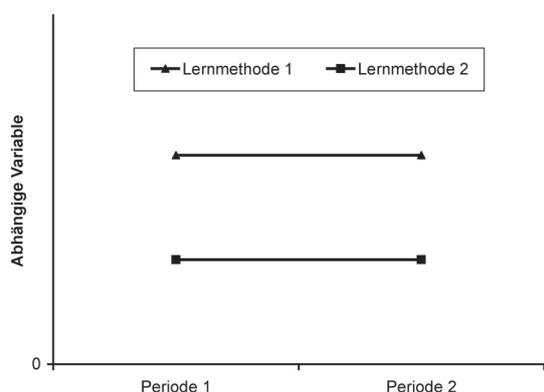


Abb. 4.8: Reiner Behandlungseffekt: Behandlungsunterschied in beiden Perioden gleich, kein Niveau-Unterschied zwischen den Perioden.

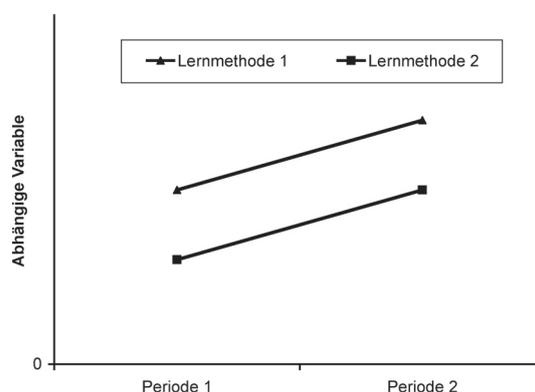


Abb. 4.9: Periodeneffekt mit Behandlungseffekt: Behandlungsunterschied in beiden Perioden gleich, Niveau-Unterschied zwischen den Perioden.

Abbildung 4.10 veranschaulicht ein Beispiel für einen **Carry-Over-Effekt**. Auf einen signifikanten Interaktionseffekt kann man auch aufgrund der nicht parallelen Linien schliessen. Bei einem Carry-Over-Effekt ist generell der Behandlungsunterschied in den beiden Perioden verschieden. In diesem Fall kann der Faktor Lernmethode nicht mehr global interpretiert werden. Stattdessen muss eine differenzierte Betrachtung der einzelnen Zellen erfolgen.

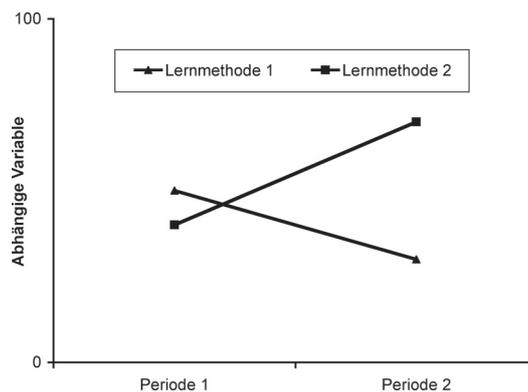


Abb. 4.10: Carry-Over-Effekt.

Signifikante Haupteffekte, sprich Behandlungs- oder Periodeneffekte, dürfen also nur global interpretiert und dabei über die Stufen des anderen Faktors hinweg generalisiert werden, wenn keine Interaktion vorliegt. Behandlungs- und Periodeneffekte können folglich nur eindeutig interpretiert werden, wenn kein Carry-Over-Effekt nachgewiesen werden kann. Die gerechneten Varianzanalysen werden deshalb zuerst immer auf einen allfälligen Interaktionseffekt hin überprüft. Besteht ein signifikanter Interaktionseffekt (Verwerfung der Nullhypothese $H_0 = \text{kein Carry-Over-Effekt}$), so wird von einem Carry-Over-Effekt ausgegangen. Man beachte, dass die Frage nach globaler oder differenzierter Interpretation von Haupteffekten nur relevant ist, wenn die Haupteffekte signifikant sind. Dieses Verfahren, bei dem zunächst ein Vortest auf einen Carry-Over-Effekt durchgeführt wird, um aufgrund der Ergebnisse weiter vorzugehen, entspricht der zweistufigen Analyseform nach Grizzle (1965). Es gibt aber ein Problem bei dieser Prozedur der Überprüfung von Carry-Over-Effekten: Das Ausschliessen eines Carry-Over-Effekts gründet darin, dass durch Nichtverwerfen der Nullhypothese im Prinzip auf deren Gültigkeit geschlossen wird, was aber streng genommen eine falsche Schlussfolgerung ist. Eine Kontrolle über die Fehlerwahrscheinlichkeit fehlt hier. Da eine fälschliche Annahme der Nullhypothese als gravierender betrachtet wird als eine irrtümliche Ablehnung, soll die Nullhypothese schon bei einer relativ hohen Irrtumswahrscheinlichkeit von 10% verworfen werden.

Bis anhin wurden signifikante Interaktionseffekte als Carry-Over-Effekte interpretiert. Diese Schlussfolgerung soll für die vorliegende Studie eingeschränkt werden. Carry-Over-Effekte sollen nur bei subjektiven Variablen wie der Beanspruchung und der emotionalen Stimmung der Upn als wahrscheinlich gelten. Bei den objektiven Variablen zur Lernwirksamkeit werden Carry-Over-Effekte nicht für möglich gehalten, da eine Beeinflussung einer (objektiven) Testleistung durch eine vorangegangene Testleistung nicht angenommen wird. Im Gegensatz dazu können subjektive Beurteilungen einer Behandlungsperiode das Ergebnis der folgenden Behandlungsperiode im Sinne eines psychologischen Carry-Over-Effekts durchaus beeinflus-

sen. Folglich werden allfällige signifikante Interaktionseffekte bei den Variablen zur Lernwirksamkeit nicht als Carry-Over-Effekte interpretiert. Interaktionseffekte zwischen Periode und Lernmethode könnten dadurch zustande kommen, dass die mit den beiden Perioden verbundenen unterschiedlichen Lerninhalte unterschiedliche Effekte der Lernmethode auf die abhängige Variable hervorrufen. Ein möglicher Interaktionseffekt könnte folglich darin gründen, dass Lerninhalt A besser mit der einen Lernmethode, Lerninhalt B hingegen besser mit der anderen Lernmethode gelernt wird. Die Gründe dafür müssten anhand der Aufgabentypen genauer analysiert werden.

4.6.2 Zusammenhänge

Die Zusammenhänge zwischen der Lernwirksamkeit und den subjektiven Variablen werden getrennt nach den beiden Lernsituationen analysiert. Diese Zusammenhänge werden mittels Rangkorrelationen nach Spearman berechnet. Zusätzlich soll mittels multipler linearer Regression versucht werden, Lernerfolg aus den Variablen Lernmethode, Periode, emotionale Stimmung und Beanspruchung vorherzusagen.

4.6.3 Subjektive Usability

Die Auswertung des SUMI-Fragebogens zur Erfassung der subjektiven Usability von AC erfolgt mit der speziellen Software SUMISCO. Anhand der Normstichprobe werden dabei die Rohwerte z-transformiert und die dazugehörigen Standardwerte so bestimmt, dass der Populationsmittelwert bei $\mu = 50$ und die Standardabweichung bei $\sigma = 10$ liegt (T-Skala). Die Standardwerte der Skalen liegen folglich in einem Wertebereich von 0 bis 100. Standardwerte ≤ 50 deuten auf eine geringe Usability hin, während Werte ≤ 40 einen deutlichen Verbesserungsbedarf der Software markieren. Gute Software erreichen Werte ≥ 60 in allen Subskalen.

5. Ergebnisse

In diesem Kapitel werden die Ergebnisse der Untersuchung beschrieben. Zu Beginn werden als Übersicht allgemeine Zusammenhänge dargestellt. Anschliessend werden die Hauptergebnisse der Untersuchung - die Lernwirksamkeit der beiden Lernmethoden - berichtet. Es folgt eine Analyse der subjektiven Variablen Beanspruchung, emotionale Stimmung und Zufriedenheit der Lernenden. Nach einer Analyse der Präferenz für ein Lernsystem und der subjektiven Usability von AC werden die Zusammenhänge zwischen Lernerfolg und den subjektiven Variablen genauer untersucht.

5.1 Allgemeine Zusammenhänge

Anhand der in Tabelle 5.1 abgebildeten Korrelationsmatrix wird eine Übersicht darüber gegeben, ob und wie die wichtigsten Variablen der Untersuchung zusammenhängen.

Tab. 5.1: Korrelationsmatrix der wichtigsten Variablen der Untersuchung.

	I	II	III	IV	V	VI	VII	VIII
I Lernmethode ^(a)	-							
II Periode ^(b)	.00	-						
III Problemlösefähigkeit (Aufgaben)	.38**	.41**	-					
IV Lernerfolg (Kurztest)	.13	-.81**	-.07	-				
V Beanspruchung (NASA)	-.13	.09	-.47**	-.31*	-			
VI Aktiviertheit (SAM)	-.15	-.01	-.28	-.12	.28	-		
VII Valenz (SAM)	.14	-.01	.43**	.07	-.61**	-.16	-	
VIII Dominanz (SAM)	.05	.04	.28	-.03	-.27	-.16	.19	-

(Korrelationen nach Spearman)

Anmerkungen: n = 48. *, **: p < .05, p < .01

^(a)Wertebereich: 1 = AC, 2 = KSM;

^(b)Wertebereich: 1 = erster Versuch (Themenbereich A), 2 = zweiter Versuch (Themenbereich B)

Die unabhängige Variable Lernmethode steht mit der abhängigen Variable Problemlösefähigkeit in einem mittelstarken Zusammenhang von $r = .38$ ($p < .01$). Dies bedeutet, dass - auf die gesamte Untersuchung bezogen - Personen, die mit dem KSM gearbeitet hatten eher eine höhere Punktzahl in den Lernaufgaben erreichten als Personen, die mit AC gearbeitet hatten. Des Weiteren zeigt sich ein mittlerer bis starker Zusammenhang von $r = .41$ ($p < .01$) zwischen der Periode und der Problemlösefähigkeit. Demzufolge wurde in der zweiten Periode zum Themenbereich B allgemein eher die höhere Punktzahl in den Lernaufgaben erreicht als in der ersten Periode zum Themenbereich A. Erstaunlicherweise verhält es sich beim Lernerfolg genau

umgekehrt: Die auffallend hohe negative Korrelation von $r = -.81$ ($p < .01$) zwischen der Periode und dem Lernerfolg besagt, dass im ersten Versuch zum Themenbereich A eher eine höhere Punktzahl im Kurztest erreicht wurde als im zweiten Versuch zum Themenbereich B. Weitere interessante Korrelationen finden sich bei den subjektiven Beurteilungen. Bei der Beanspruchung zeigt sich ein starker negativer Zusammenhang von $r = -.47$ mit der Problemlösefähigkeit ($p < .01$). Je höher also die subjektive Beanspruchung beim Arbeiten mit den Lernsystemen war, desto weniger Punkte wurden in den Lernaufgaben erreicht. Die Beanspruchung wirkt sich hier negativ auf die Problemlösefähigkeit der Upn aus. Derselbe Effekt etwas geringer ausgeprägt zeigt sich mit einer negativen Korrelation von $r = -.31$ ($p < .05$) beim Lernerfolg. Ein mittlerer bis starker positiver Zusammenhang von $r = .43$ ($p < .01$) ist zwischen Valenz und Problemlösefähigkeit ersichtlich. Dies bedeutet, dass je höher die erreichte Punktzahl in den Lernaufgaben war, desto positivere Gefühle wurden dadurch ausgelöst.

Aufgrund der hohen Korrelationen der Lernwirksamkeit mit der Periode, sollen im Folgenden die Zusammenhänge zwischen der Lernmethode und den Lernvariablen *Lernerfolg* und *Problemlösefähigkeit* getrennt nach den beiden Perioden betrachtet werden (Tabellen 5.2 und 5.3).

Tab. 5.2: Korrelationsmatrix nur von der ersten Periode (Thema A).

	I	II	III
I Lernmethode ^(a)	-		
II Problemlösefähigkeit (Lernaufgaben)	.07	-	-
III Lernerfolg (Kurztest)	.14	.40	

(Korrelationen nach Spearman)

Anmerkungen: $n=24$.

^(a)Wertebereich: 1=AC, 2=KSM;

Tab. 5.3: Korrelationsmatrix nur von der zweiten Periode (Thema B).

	I	II	III
I Lernmethode ^(a)	-		
II Problemlösefähigkeit (Lernaufgaben)	.63**	-	-
III Lernerfolg (Kurztest)	.36	.61**	

(Korrelationen nach Spearman)

Anmerkungen: $n=24$. **: $p < .01$

^(a)Wertebereich: 1=AC, 2=KSM;

Während in der ersten Periode zum Themenbereich A kaum oder nur schwache Korrelationen der Lernmethode mit der Problemlösefähigkeit und mit dem Lernerfolg bestehen ($r = .07$ bzw.

$r = .14, p > .05$) (Tabelle 5.2), so fallen diese Korrelationen in der zweiten Periode deutlich stärker aus (Tabelle 5.3). Der starke und zudem sehr signifikante Zusammenhang von $r = .63$ zwischen Lernmethode und Problemlösefähigkeit ($p < .01$) weist darauf hin, dass Personen, die in der zweiten Periode mit dem KSM gearbeitet hatten eher eine höhere Punktzahl in den Lernaufgaben erreichten, als Personen, die mit AC gearbeitet hatten. Auch beim Lernerfolg zeigen sich diese Zusammenhänge anhand einer mittleren bis starken Korrelation mit der Lernmethode ($r = .36, p > .05$). Dieser Sachverhalt, dass je nach Periode sich solch unterschiedliche Resultate bei der Lernwirksamkeit ergeben, wird in den weiteren Analysen berücksichtigt.

5.2 Lernwirksamkeit

In den folgenden Kapiteln wird die Lernwirksamkeit der beiden Lernmethoden AC und KSM genauer analysiert. In einem ersten Schritt werden die Verteilungen der Leistung in den Lernaufgaben und im Kurztest beschrieben, um danach die Unterschiede in der praktischen Problemlösefähigkeit und im Lernerfolg zwischen den beiden Lernmethoden genauer zu untersuchen.

5.2.1 Verteilung der Leistung in den Lernaufgaben und im Kurztest

Als Überblick wird zuerst die Häufigkeitsverteilung der allgemeinen Leistung in den Lernaufgaben und im Kurztest über beide Perioden hinweg beschrieben. Danach werden die Verteilungen der beiden Lernmethoden vergleichend dargestellt, sowohl für beide Perioden gleichzeitig als auch getrennt nach Periode 1 und Periode 2.

Verteilung der Leistung in den Lernaufgaben allgemein

Die Punktzahlen in den Lernaufgaben reichten in der untersuchten Stichprobe von 53 bis 100 Punkte, wobei eine Person gar keine Punkte erreichte (Abbildung 5.1). Der Mittelwert liegt bei 86.12 Punkten, die Standardabweichung beträgt 17.99. Nach Tabachnik und Fidell (1996, S. 73) kann man von einer Normalverteilung sprechen, wenn die Werte der Schiefe (skewness) und Steilheit (kurtosis) gleich Null sind. Mit einer Schiefe von -2.60 handelt es sich im vorliegenden Fall um eine stark linksschiefe Verteilung. Da dieser Wert signifikant von Null abweicht, kann nicht von einer Normalverteilung ausgegangen werden, wie auch aus Abbildung 5.1 ersichtlich ist. Abgesehen von der Up mit 0 Punkten erreichten alle Upn eine Punktzahl über dem arithmetischen Mittel von 50 Punkten. Die Lernenden waren demzufolge alle erfolgreich bis sehr erfolgreich in den Lernaufgaben, was auf eine gute bis sehr gute praktische Problemlöseleistung mit beiden Lernsystemen hinweist.

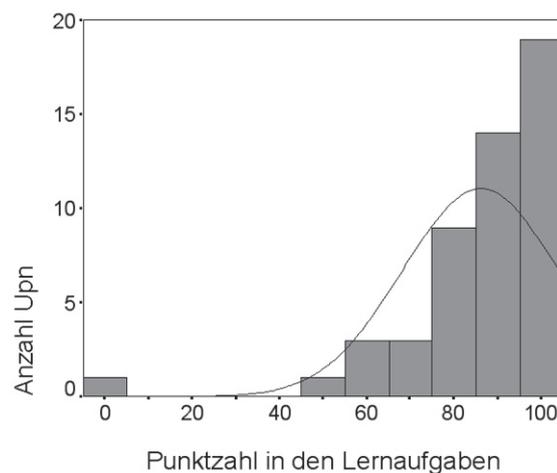


Abb. 5.1: Häufigkeitsverteilung der Leistung in den Lernaufgaben über beide Perioden hinweg ($M = 86.12$, $SD = 17.99$, $N = 50$).

Nach Tabachnik und Fidell (1996) sind z-Werte > 3.29 potenzielle Ausreisser. In den vorliegenden Daten trifft das Kriterium auf einen Wert zu: er hat eine Ausprägung von 0 Punkten (z-Wert: -4.79). Um Verzerrungen zu vermeiden, sollte dieser Wert angepasst oder ausgeschlossen werden (Tabachnik & Fidell, 1996). Im vorliegenden Fall werden alle Werte dieser Up aus den weiteren Analysen ausgeschlossen. Aufgrund der Eliminierung dieses Ausreissers und wegen einer Up, welche während der gesamten Untersuchung abwesend war, reduziert sich die Stichprobe auf $N = 24$.

Verteilung der Leistung im Kurztest allgemein

Im Folgenden wird bereits auf die vom Ausreisser befreiten Daten Bezug genommen. In den beiden Kurztests wurden insgesamt Werte zwischen 6 und 100 Punkte erreicht. Der Mittelwert liegt bei 59.47 Punkten, die Standardabweichung beträgt 26.66. Die Schiefe dieser Verteilung hat eine Ausprägung von -0.12 und die Steilheit beträgt -1.15 . Da beide Werte nicht signifikant von Null abweichen, kann hier von einer annähernden Normalverteilung ausgegangen werden (Tabachnik & Fidell, 1996). Das Histogramm in Abbildung 5.2 zeigt zwar keine eindeutige eingipflige Verteilung, bei den folgenden Auswertungen wird aber dennoch von einer annähernden Normalverteilung des Konstrukts Lernerfolg ausgegangen.

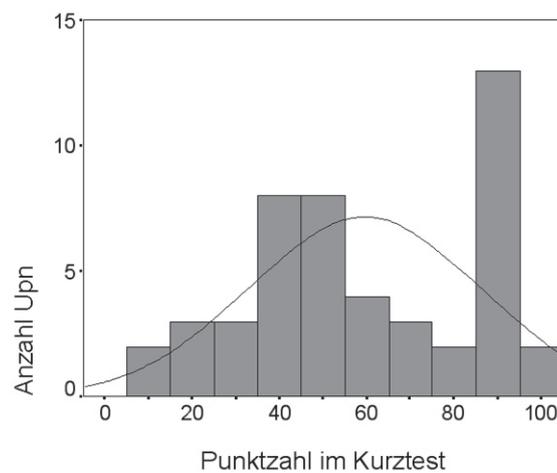


Abb. 5.2: Häufigkeitsverteilung der Leistung im Kurztest über beide Perioden hinweg ($M = 59.47$, $SD = 26.66$, $N = 48$).

Verteilungen der Leistung in den Lernaufgaben pro Lernmethode

Im Folgenden werden die Verteilungen der Leistung in den Lernaufgaben getrennt für Personen, die mit AC, und für jene, die mit dem KSM gearbeitet haben, dargestellt. Gemäss dem vorliegenden Untersuchungsdesign gründen beide Verteilungen auf derselben Stichprobe. Die Verteilungen sind hier noch nicht vom Ausreisser bereinigt, dies deshalb um genauer zu untersuchen, mit welchem Lernsystem und in welcher Periode (siehe später, Abbildung 5.5) dieser extreme Wert zu Stande gekommen ist. Für die späteren varianzanalytischen Berechnungen wird der Ausreisser aber eliminiert.

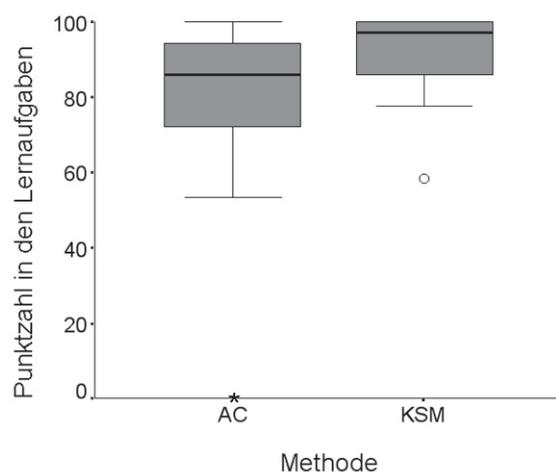


Abb. 5.3: Verteilung der Leistung in den Lernaufgaben getrennt nach AC und dem KSM ($N_{AC} = 25$; $N_{KSM} = 25$).

Die Abbildung 5.3 veranschaulicht die Verteilungen der Leistung in den Lernaufgaben gruppiert nach den beiden Lernmethoden AC und KSM. Die Verteilungen beziehen sich (wie auch jene in Abbildung 5.4) auf die gesamte Untersuchung, also auf beide Perioden, was bedeutet, dass der Vergleich der beiden Lernsysteme sozusagen “innerhalb” der Upn erfolgt. Die Verteilungen werden mit Boxplots visualisiert.

Die Verteilungen in Abbildung 5.3 verdeutlichen, dass die Lernaufgaben mit beiden Lernsystemen überdurchschnittlich gut gelöst wurden. Mit dem KSM wurde aber im Schnitt eine höhere Punktzahl erreicht als mit AC, was sich in einem höheren Median zeigt ($Md_{KSM} = 97.22$ vs. $Md_{AC} = 86.11$). Was den Ausreisser aus Abbildung 5.1 betrifft, so sieht man in Abbildung 5.3, dass dieser Wert beim Arbeiten mit AC zu Stande gekommen ist. Dies bedeutet, dass eine Person nicht in der Lage war, die Lernaufgaben mit AC zu lösen. Wie bereits erwähnt, wird diese Up aus den folgenden statistischen Analysen ausgeschlossen. Während bei AC die Werte neben diesem Extremwert von 0 Punkten zwischen 53 und 100 Punkten variieren, bewegen sich beim KSM die Leistungen - abgesehen von einem Ausreisser bei 58 Punkten - im oberen Bereich zwischen 78 und 100 Punkten. Eine genauere Analyse der Daten zeigt, dass der Extremwert bei AC und der Ausreisser beim KSM von derselben Person stammen. Dies rechtfertigt den Ausschluss dieser Up aus der gesamten Analyse der Daten. Die Streuung der Wertebereiche ist beim KSM kleiner als bei AC, und es handelt sich sowohl bei AC als auch beim KSM um schiefe Verteilungen mit extremen Werten im unteren Bereich.

Verteilungen der Leistung im Kurztest pro Lernmethode

Folgend werden die Verteilungen der Leistung im Kurztest gruppiert nach den beiden Lernmethoden beschrieben. Die Up, welche aufgrund des Ausreissers ausgeschlossen wird (s. oben) ist in dieser Stichprobe bereits nicht mehr enthalten. Dies erklärt das N von 24.

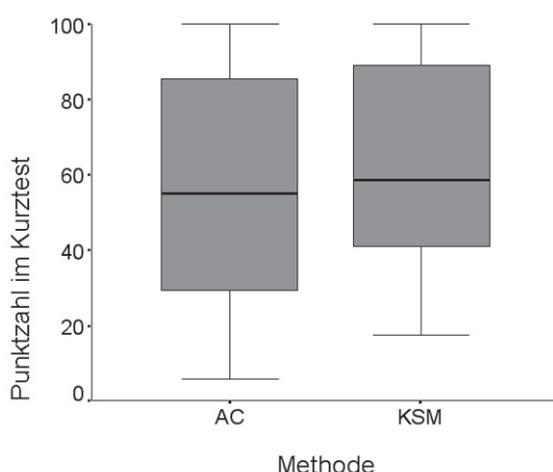


Abb. 5.4: Verteilung der Leistung im Kurztest getrennt nach AC und dem KSM ($N_{AC}=24$; $N_{KSM}=24$).

Die beiden Verteilungen fallen sehr ähnlich aus (Abbildung 5.4). Es zeigt sich aber, dass Personen, die mit dem KSM gelernt hatten im Schnitt eine leicht höhere Punktzahl im Kurztest erreichten als Personen, die den Stoff mit AC erarbeitet hatten ($Md_{KSM} = 63.48$; $Md_{AC} = 55.04$). Beide Mediane sind nicht weit vom arithmetischen Mittel von 50 Punkten entfernt. Bei AC streuen die Werte aber etwas stärker als beim KSM. Die Verteilungen variieren zwischen einem Wertebereich von 6 und 100 Punkten bei AC und zwischen 18 und 100 Punkten beim KSM. Beide Verteilungen sind annähernd symmetrisch.

Verteilungen der Leistung in den Lernaufgaben pro Lernmethode und Periode

Im Folgenden werden getrennt für die beiden Perioden die Verteilungen der erreichten Punktzahl in den Lernaufgaben für beide Lernmethoden vergleichend dargestellt. Es folgt zuerst in Abbildung 5.5 die Darstellung mit dem Ausreisser, zur Analyse in welcher Periode dieser Wert entstanden ist. Die Verteilungen der vom Ausreisser bereinigten Daten sind in Abbildung 5.6 ersichtlich.

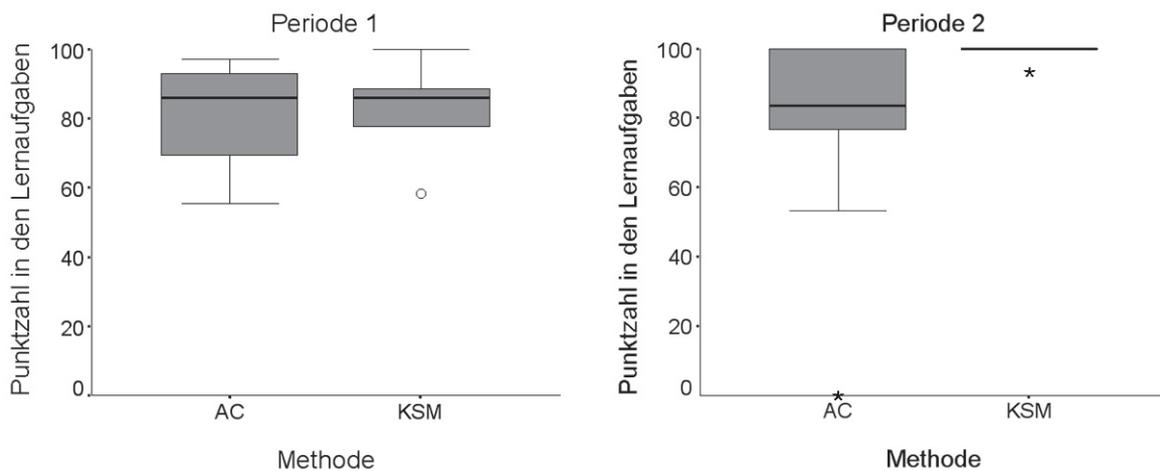


Abb. 5.5: Verteilung der erreichten Punktzahl in den Lernaufgaben getrennt nach AC und dem KSM und separat für beide Perioden (mit Ausreisser!).

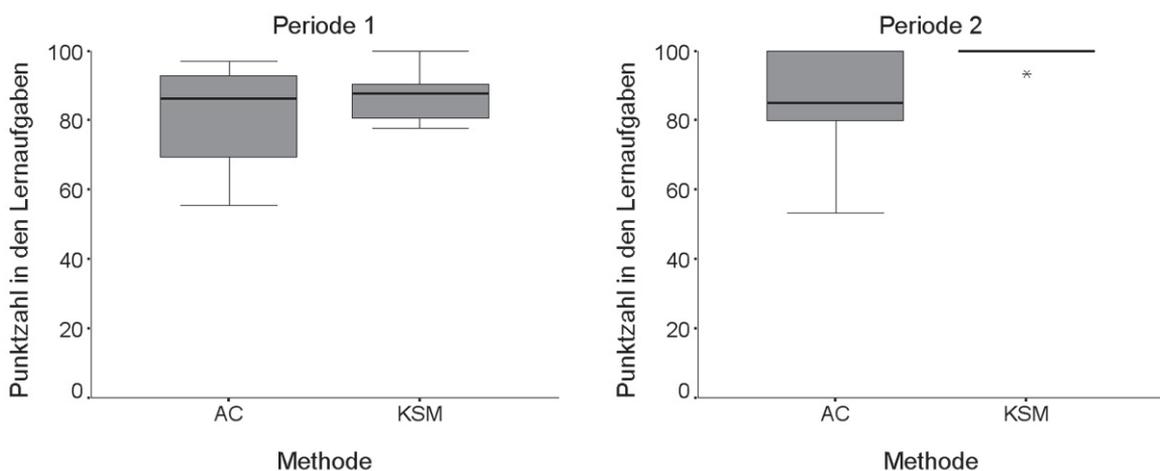


Abb. 5.6: Verteilung der erreichten Punktzahl in den Lernaufgaben getrennt nach AC und dem KSM und separat für beide Perioden (ohne Ausreisser!).

Abbildung 5.5 zeigt, dass der Extremwert von 0 Punkten mit AC in der zweiten Periode zu Stande gekommen ist. Als Überblick werden die wichtigsten Kennwerte aus Abbildung 5.6 tabellarisch dargestellt (Tabelle 5.4). Die dunkelgrau eingefärbten Zahlen entstammen der Gruppe 1, also denjenigen Personen, die zuerst mit AC, dann mit dem KSM gelernt haben. Die hellgrau eingefärbten Zahlen stammen von der Gruppe 2, d.h. von denjenigen Personen, die zuerst mit dem KSM dann mit AC gelernt haben.

Tab. 5.4: Übersicht über die wichtigsten Kennwerte aus Abbildung 5.6.

	Periode 1 (Themenbereich A)				Periode 2 (Themenbereich B)			
	Q ₁	Md	Q ₃	N	Q ₁	Md	Q ₃	N
AC	68.06	86.11	93.75	12	78.33	85.00	100	12
KSM	79.17	87.50	90.97	12	100	100	100	12

Anmerkungen: Q₁: 1. Quartil; Q₃: 3. Quartil; (Interquartilbereich $d_Q = Q_3 - Q_1$)

Dunkelgrau hinterlegte Zahlen entstammen der Gruppe 1 (AC -> KSM); hellgrau hinterlegte Zahlen entstammen der Gruppe 2 (KSM -> AC).

Bei der ersten Periode unterscheidet sich der Median der erreichten Punktzahl in den Lernaufgaben mit AC und mit dem KSM nur wenig ($Md_{AC} = 86.11$; $Md_{KSM} = 87.50$). Im Schnitt lösten demnach beide Lerngruppen die Aufgaben etwa gleich gut. Die Werte streuen bei AC aber stärker als beim KSM, was am grösseren Interquartilbereich (d_Q) ersichtlich ist ($d_Q = 25.69$ bei AC vs. $d_Q = 11.80$ beim KSM). Mit AC wurden im Vergleich zum KSM überproportional häufiger tiefere Werte in den Lernaufgaben erreicht. Während bei AC 25% aller Werte zwischen 56 und 68 Punkten liegen, so sind beim KSM unterhalb von 78 Punkten gar keine Werte enthalten. Zudem wurde das Maximum von 100 Punkten mit dem KSM nicht aber mit AC erreicht. Mit AC wurde eine maximale Punktzahl von 97 Punkten erzielt.

Bei der zweiten Periode zeigt sich ein beachtlicher Unterschied zwischen den beiden Lerngruppen. Im Schnitt wurden die Aufgaben mit dem KSM erheblich besser gelöst ($Md = 100$) als mit AC ($Md = 85$). Beim KSM zeigt sich aber keinerlei Variation in den Werten. Alle Upn, ausser einer Person mit 93 Punkten, erreichten die maximale Punktzahl von 100 Punkten. Bei AC variieren die Werte hingegen deutlich zwischen einem Minimum von 53 Punkten und einem Maximum von 100 Punkten.

Vergleicht man die beiden Perioden miteinander, so zeigt sich, dass mit AC in beiden Perioden im Schnitt vergleichbare Werte erreicht wurden ($Md = 86.11$ vs. $Md = 85.00$). Mit dem KSM hingegen wurden in der zweiten im Vergleich zur ersten Periode im Schnitt deutlich bessere Resultate erzielt ($Md = 87.50$ vs. $Md = 100$).

Zusammengefasst lässt sich sagen, dass in beiden Perioden die Werte in den Lernaufgaben bei AC stärker streuen als beim KSM, und mit AC häufiger tiefere Werte erreicht wurden als mit dem KSM. Während sich die Leistungen in den Lernaufgaben mit AC und dem KSM bei der ersten Periode im Schnitt kaum voneinander unterscheiden, zeigen sich bei der zweiten Periode deutliche Unterschiede zu Gunsten des KSM, wobei ein Median von 100 Punkten beim KSM einem extremen Resultat gleichkommt.

Verteilungen der Leistung im Kurztest pro Lernmethode und Periode

In Abbildung 5.7 sind die Verteilungen für die Resultate im Kurztest vergleichend für AC und das KSM und separat nach den beiden Perioden dargestellt.

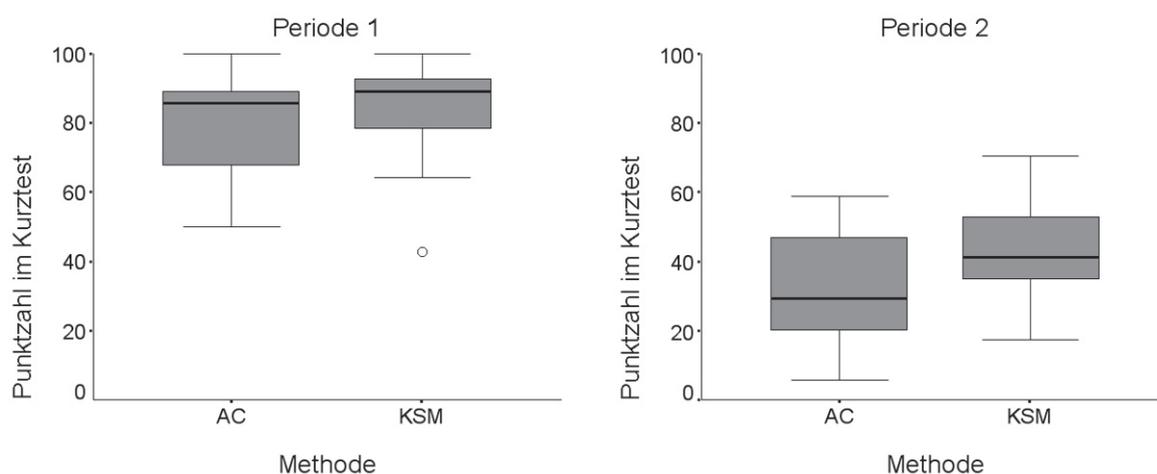


Abb. 5.7: Verteilung der erreichten Punktzahl im Kurztest getrennt nach AC und dem KSM und separat für beide Perioden.

Die wichtigsten Verteilungskennwerte aus Abbildung 5.7 sind in Tabelle 5.5 dargestellt.

Tab. 5.5: Übersicht über die wichtigsten Kennwerte aus Abbildung 5.7.

	Periode 1 (Themenbereich A)				Periode 2 (Themenbereich B)			
	Q ₁	Md	Q ₃	N	Q ₁	Md	Q ₃	N
AC	66.07	85.71	91.07	12	19.12	29.41	47.06	12
KSM	78.57	89.29	92.86	12	35.29	41.18	52.94	12

Anmerkungen: Q₁: 1. Quartil; Q₃: 3. Quartil; (Interquartilbereich $d_Q = Q_3 - Q_1$)

Dunkelgrau hinterlegte Zahlen entstammen der Gruppe 1 (AC-> KSM); hellgrau hinterlegte Zahlen entstammen der Gruppe 2 (KSM -> AC).

Beim Kurztest der ersten Periode zeigen sich sehr ähnliche Verteilungen der beiden Lerngruppen. Personen, die mit AC gelernt hatten, unterscheiden sich in ihrer Leistung im Schnitt kaum von Personen, die mit dem KSM gelernt hatten ($Md_{AC} = 85.71$; $Md_{KSM} = 89.29$). Die Werte streuen bei AC aber stärker als beim KSM ($d_Q = 25$ bei AC vs. $d_Q = 14.29$ beim KSM). Es handelt sich bei beiden Methoden um nahezu symmetrische Verteilungen, welche leicht linksschief sind. Beim KSM ist ein Ausreisser bei 43 Punkten zu verzeichnen. Bei AC streuen die Werte unterhalb des Medians stärker als beim KSM.

In der zweiten Periode wurden im Vergleich zur ersten Periode insgesamt deutlich niedrigere Punktzahlen im Kurztest erreicht. Mit einem Median von 41.18 waren hier Personen, die mit dem KSM gelernt hatten im Schnitt deutlich besser als Personen, die mit AC gelernt hatten ($Md = 29.41$). Wiederum streuen die Werte bei AC stärker als beim KSM ($d_Q = 27.94$ bei AC vs. $d_Q = 17.65$ beim KSM). Bei beiden Lerngruppen handelt es sich um annähernd symmetrische Verteilungen, im Gegensatz zur ersten Periode sind sie hier aber leicht rechtsschief.

Zusammengefasst zeigt sich, dass sich in beiden Perioden die Verteilungen der Leistung im Kurztest der beiden Lerngruppen nur wenig voneinander unterscheiden. Grössere Unterschiede finden sich in der zweiten Periode. Des Weiteren waren die Leistungen im Kurztest bei beiden Lernmethoden in der zweiten Periode deutlich schlechter als in der ersten Periode.

5.2.2 Unterschiede in der Lernwirksamkeit zwischen den beiden Lernmethoden Augmented Chemistry und Kugel-Stab-Modell

In den folgenden Kapiteln wird die Hauptfragestellung der vorliegenden Studie untersucht, inwiefern sich die Lernwirksamkeit der beiden Lernsysteme AC und KSM voneinander unterscheidet. Die beiden Variablen *praktische Problemlösefähigkeit* und *Lernerfolg* werden mit der zweifaktoriellen Varianzanalyse auf einen Behandlungseffekt hin analysiert.

Praktische Problemlösefähigkeit

In diesem Kapitel werden die Leistungsmittelwerte in den Lernaufgaben von Personen, die mit AC, und solchen, die mit dem KSM gelernt haben, miteinander verglichen. Die Qualität der bearbeiteten Lernaufgaben soll aufzeigen, wie gut die praktische Problemlösefähigkeit mit den beiden Lernmethoden war, und folglich wie gut die Benutzer mit dem jeweiligen Lernsystem zurechtkamen. In Tabelle 5.6 sind die Mittelwerte der Leistung in den Lernaufgaben für Personen, die mit AC, und Personen, die mit dem KSM gearbeitet haben und separat für beide Perioden abgebildet.

Tab. 5.6: Mittlere Leistung in den Lernaufgaben mit AC und dem KSM getrennt nach der Periode.

	Problemlösefähigkeit Periode 1			Problemlösefähigkeit Periode 2			Total (ungewichtetes Mittel)		
	M	SD	N	M	SD	N	M	SD	N
AC	81.94	14.48	12	85.56	14.73	12	83.75	14.40	24
KSM	87.04	7.24	12	99.44	1.93	12	93.22	8.18	24
Total	84.49	11.49	24	92.50	12.48	24			

Anmerkung: Die dunkelgrau eingefärbten Zahlen entstammen der Gruppe 1 (AC -> KSM), die hellgrau eingefärbten Zahlen der Gruppe 2 (KSM -> AC). Die übrigen Zahlen beziehen sich auf beide Gruppen gleichzeitig (within-subjects).

Die stark linksschiefe Verteilung der praktischen Problemlösefähigkeit (siehe Abbildung 5.1) führt dazu, dass die beabsichtigte Varianzanalyse - die ja auf der Normalverteilung basiert - nicht mehr ohne weiteres anwendbar ist. Um eine approximative symmetrische Verteilung zu erhalten, wurden die Daten transformiert, und zwar wurde nach mehreren Versuchen die quadratische Transformation als die optimalste angesehen. Durch den Ausschluss des Ausreißers und durch die Transformation konnte die stark negative Schiefe von -2.60 auf -0.86 reduziert werden. Die Voraussetzung homogener Stichprobenvarianzen für eine Varianzanalyse konnte nicht erfüllt werden. Die transformierte Verteilung ist in Abbildung 5.8 dargestellt.

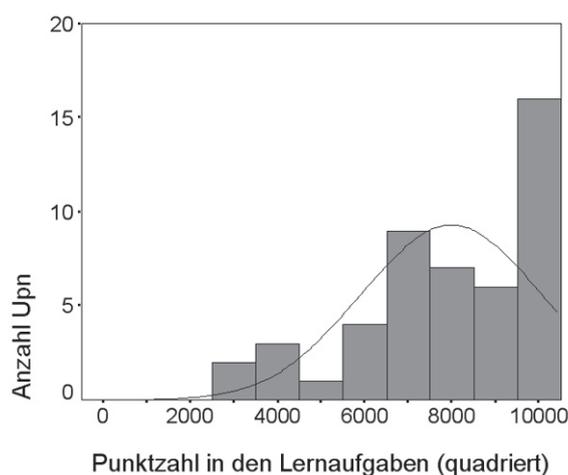


Abb. 5.8: Transformierte Verteilung der Leistung in den Lernaufgaben über beide Perioden hinweg ($N = 48$).

Die zweifaktorielle Varianzanalyse mit den transformierten Daten zeigt, dass der Interaktionseffekt der beiden Faktoren Lernmethode und Periode knapp nicht signifikant ausfällt ($F = 2.67$, $df = 1/44$, $p = .11$). Folglich können die Haupteffekte jeweils global interpretiert und dabei über die Stufen des anderen Faktors hinweg generalisiert werden. Für die Abbildung der Natur der Effekte aus der Varianzanalyse wird das Interaktionsdiagramm für den Faktor Periode gewählt (Abbildung 5.9).

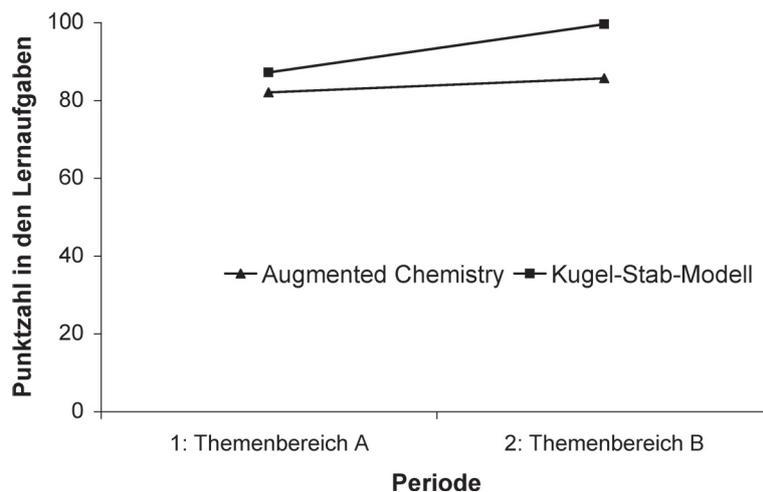


Abb. 5.9: Die durchschnittliche Leistung in den Lernaufgaben (praktische Problemlösefähigkeit) pro Lernmethode und nach der Periode.

In beiden Perioden wurde mit dem KSM eine durchschnittlich bessere Leistung in den Lernaufgaben erbracht. Der Leistungsunterschied zwischen den beiden Lernmethoden ist in der zweiten Periode stärker ausgeprägt als in der ersten Periode. Tabelle 5.6 macht deutlich, dass - global gesehen - mit dem KSM durchschnittlich eine höhere Punktzahl in den Lernaufgaben erreicht wurde ($M = 93.22$) als mit AC ($M = 83.75$). Wie der in der Varianzanalyse geprüfte Haupteffekt der Lernmethode zeigt, ist dieser Unterschied signifikant ($F = 9.27$, $df = 1/44$, $p = .004$). Nach Cohen (1988; zit. nach Hussy & Jain, 2002, S. 157) handelt es sich mit $d = .76$ um einen grossen und somit nicht unbedeutenden Effekt der Lernmethode auf die praktische Problemlösefähigkeit. Die Aussage, dass Personen, die mit dem KSM arbeiten über eine bessere praktische Problemlösefähigkeit verfügen als Personen, die mit AC arbeiten, kann somit verallgemeinert werden, und der Effekt ist praktisch bedeutsam. Der geprüfte Haupteffekt der Periode ist ebenfalls signifikant ($F = 8.06$, $df = 1/44$, $p = .01$) und kann wiederum global und über beide Lernmethoden hinweg interpretiert werden. Folglich zeigten die untersuchten Personen in der zweiten Periode (unabhängig von der Lernmethode) mit einem durchschnittlich besseren Ergebnis in den Lernaufgaben ($M = 92.50$) die bessere praktische Problemlösefähigkeit als in der ersten Periode ($M = 84.49$).

Kurz zusammengefasst können folgende Aussagen gemacht werden: Die praktische Problemlösefähigkeit war mit dem KSM signifikant besser als mit AC. Des Weiteren zeigten die Lernenden in der zweiten Periode allgemein die bessere praktische Problemlösefähigkeit als in der ersten Periode. Zu berücksichtigen ist aber, dass wichtige Voraussetzungen der Varianzanalyse wie Normalverteilung und homogene Stichprobenvarianzen nicht erfüllt werden konnten.

Lernerfolg

Im Folgenden wird analysiert, inwiefern sich Personen, die mit AC gelernt haben, von solchen, die mit dem KSM gelernt haben, in ihrem Lernerfolg unterscheiden. Hierbei werden die Leistungsmittelwerte im Kurztest von Personen, die mit AC, und von solchen, die mit dem KSM gelernt haben, miteinander verglichen. Es wird von annähernd normalverteilten Daten ausgegangen (siehe Abbildung 5.4).

Tab. 5.7: Mittlere Leistung in den Kurztests mit AC und dem KSM getrennt nach der Periode.

	Lernerfolg Periode 1 (Thema A)			Lernerfolg Periode 2 (Thema B)			Total (ungewichtetes Mittel)		
	M	SD	N	M	SD	N	M	SD	N
AC	78.57	15.53	12	32.35	16.54	12	55.46	28.35	24
KSM	83.33	15.92	12	43.63	13.14	12	63.48	24.80	24
Total	80.95	15.57	24	37.99	15.71	24			

Anmerkung: Die dunkelgrau eingefärbten Zahlen entstammen der Gruppe 1 (AC -> KSM), die hellgrau eingefärbten Zahlen der Gruppe 2 (KSM -> AC). Die übrigen Zahlen beziehen sich auf beide Gruppen gleichzeitig (within-subjects).

Tabelle 5.7 gibt einen Überblick über die Mittelwerte in den Kurztests der beiden Perioden jeweils getrennt nach Personen, die mit AC, und jenen, die mit dem KSM gelernt haben.

Die zweifaktorielle Varianzanalyse ergibt, dass der Interaktionseffekt der beiden Faktoren Lernmethode und Periode nicht signifikant ist ($F = 0.54$, $df = 1/44$, $p = .47$). Die Natur der Effekte aus der Varianzanalyse ist in Abbildung 5.10 mit dem Interaktionsdiagramm für den Faktor Periode dargestellt.

Der Leistungsunterschied zwischen den beiden Lernmethoden ist in der zweiten Periode etwas ausgeprägter als in der ersten Periode. Bei diesem unterschiedlichen Einfluss der Lernmethode je nach Periode kann aber aufgrund der nicht signifikanten Interaktion der Zufall nicht ausgeschlossen werden.

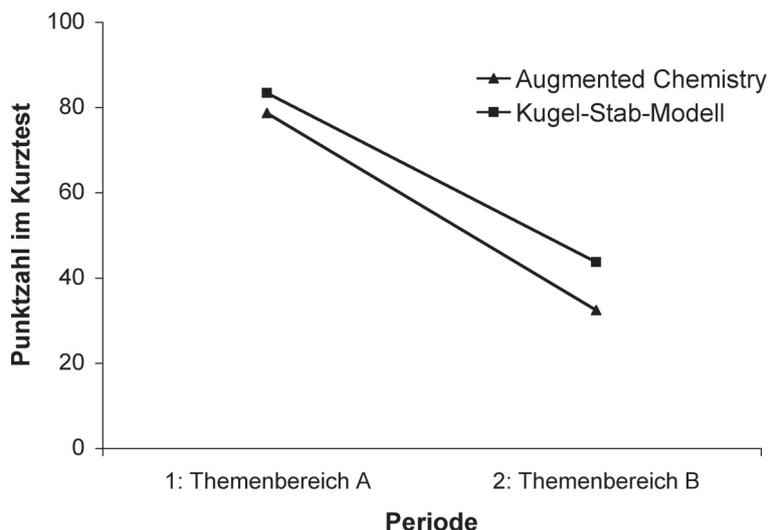


Abb. 5.10: Die durchschnittliche Leistung im Kurztest (Lernerfolg) pro Lernmethode und nach der Periode.

Wie in Tabelle 5.7 zu sehen ist, erreichten Personen - unabhängig von der Periode -, welche mit dem KSM gelernt hatten durchschnittlich eine höhere Punktzahl im Kurztest ($M = 63.48$) als Personen, die mit AC gelernt hatten ($M = 55.46$). Dieser Unterschied ist eher klein, wenn man berücksichtigt, dass die Punkteskala von 0 bis 100 reicht. Der geprüfte Haupteffekt der Lernmethode ist knapp nicht signifikant ($F = 3.28$, $df = 1/44$, $p = .08$) und der Effekt ist mit $d = .30$ klein. Daraus kann geschlossen werden, dass die Lernmethode einen kleinen, aber nicht generalisierbaren Zusammenhang mit dem Lernerfolg hat. Der geprüfte Haupteffekt der Periode ist dagegen hoch signifikant ($F = 94.14$, $df = 1/44$, $p < .001$) und widerspiegelt die Tatsache, dass in der zweiten Periode ($M = 37.99$), unabhängig von der Lernmethode, durchschnittlich sehr viel niedrigere Werte im Kurztest erreicht wurden als in der ersten Periode ($M = 80.95$).

Es zeigt sich somit nicht der erhoffte Effekt, dass der Lernerfolg mit AC besser ausfällt als mit dem herkömmlichen KSM. Mit dem KSM wurden sogar leicht bessere Resultate erzielt. Die Lernmethode steht aber mit dem Lernerfolg in keinem direkten Zusammenhang, der nicht auch durch den Zufall erklärt werden könnte. Daraus kann man zumindest schliessen, dass die Lernziele, die bisher mit dem KSM erreicht wurden, auch mit AC erarbeitet werden können.

5.3 Beanspruchung

In den folgenden Kapiteln wird die von den Lernenden empfundene Gesamtbeanspruchung während der Lernsituation genauer analysiert. Das Ziel liegt darin, herauszufinden, ob und inwiefern sich Personen, die mit AC gelernt haben, von solchen, die mit dem KSM gelernt haben, in ihrer Gesamtbeanspruchung sowie auch in Teilbereichen der Beanspruchung voneinander unterscheiden.

Vorerst werden als Überblick die Häufigkeitsverteilungen der Upn auf die Gesamtbeanspruchung graphisch dargestellt. Dabei werden neben einer Gesamtdarstellung (Abbildung 5.11) die Verteilungen auch getrennt für die beiden Lernmethoden visualisiert (Abbildung 5.12). Da die subjektiven Variablen Beanspruchung und emotionale Stimmung (Aktiviertheit, Valenz, Dominanz) nur sehr schwach mit der Periode korrelieren (siehe Tabelle 5.1), werden an dieser Stelle die Verteilungen nicht wie bei den Lernvariablen (Problemlösefähigkeit und Lernerfolg) getrennt nach Periode analysiert.

5.3.1 Verteilung der Untersuchungspersonen hinsichtlich ihrer Gesamtbeanspruchung

In Abbildung 5.11 werden die Beanspruchungswerte über beide Perioden hinweg zusammengefasst. Dies bedeutet, dass in der Grafik von jeder Up zwei Urteile enthalten sind, eines von der ersten und eines von der zweiten Periode. Das erklärt auch die Anzahl an $N = 46$ Upn.

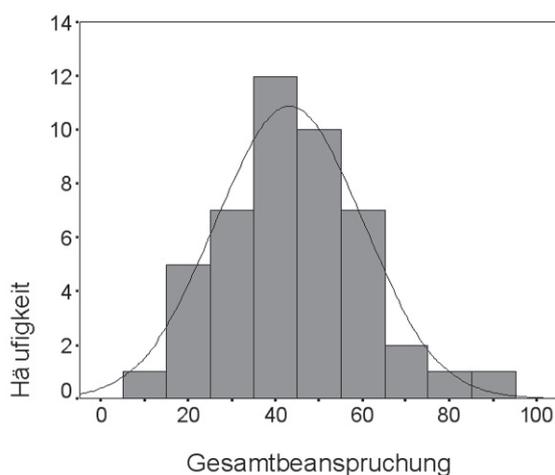


Abb. 5.11: Häufigkeitsverteilung der Gesamtbeanspruchung bei den Upn über beide Perioden hinweg ($M=43.20$, $SD=16.80$, $N=46$).

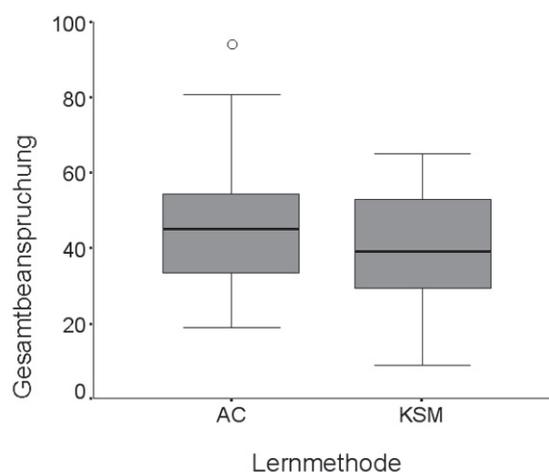


Abb. 5.12: Verteilung der Gesamtbeanspruchung getrennt nach AC und dem KSM ($N_{AC} = N_{KSM} = 23$).

Wie Abbildung 5.11 zeigt, handelt es sich um annähernd normalverteilte Daten (Schiefe = .57, Steilheit = .87). Durchschnittlich fühlten sich die Upn mässig belastet, was aus dem Mittelwert von $M = 43.20$ hervorgeht. Dieser Mittelwert befindet sich leicht unter dem arithmetischen Mittel von 50 Punkten.

Vergleicht man die Verteilung der Beanspruchung von Personen, die mit AC gelernt haben, mit jener von Personen, die mit dem KSM gelernt haben (Abbildung 5.12), so zeigen sich nur geringe Unterschiede. Beim Lernen mit AC traten aber im Gegensatz zum Lernen mit dem KSM auch überaus hohe Beanspruchungswerte auf: Bei AC wurde ein Maximum von 94 Punkten erreicht, während beim KSM das Maximum bei lediglich 65 Punkten lag. Extrem hohe Beanspruchungen kamen folglich nur beim Lernen mit AC vor, nicht aber beim Lernen mit dem KSM. Im Gegensatz dazu traten extrem tiefe Beanspruchungswerte eher beim Lernen mit dem KSM auf: Beim KSM wurde ein Minimum von 9 Punkten, bei AC ein Minimum von 19 Punkten erreicht.

5.3.2 Unterschiede in der Beanspruchung zwischen Personen, die mit Augmented Chemistry, und solchen, die mit dem Kugel-Stab-Modell gelernt haben

Im Folgenden wird die Beanspruchungsstruktur von Personen, die mit AC, und solchen, die mit dem KSM gelernt haben, genauer untersucht. Dabei wird zuerst auf die Gesamtbeanspruchung (gewichteter Gesamtwert) eingegangen, um anschliessend die sechs Subskalen der Beanspruchung einzeln zu analysieren. Durch Mittelwertvergleiche der Beanspruchungswerte werden die beiden Lernmethoden auf mögliche Unterschiede hin untersucht. Allfällige Unterschiede werden wiederum anhand des varianzanalytischen Designs über den Haupteffekt der Lernmethode analysiert. Mit der vorgängigen Testung nach Interaktionseffekten zwischen den Faktoren Lernmethode und Periode, werden mögliche Carry-Over-Effekte aufgedeckt.

Gesamtbeanspruchung

Zur Erinnerung: Für die Gesamtbeanspruchung wurden die Ratings in den einzelnen Subskalen mit ihrer Gewichtung verrechnet. In Tabelle 5.8 sind die Mittelwerte in der Gesamtbeanspruchung für Personen, die mit AC, und Personen, die mit dem KSM gearbeitet haben, und separat für beide Perioden abgebildet.

Während für die Variable der Beanspruchung von annähernd normalverteilten Daten ausgegangen wird (Abbildung 5.12), so gilt die Voraussetzung homogener Stichprobenvarianzen für eine Varianzanalyse als nicht erfüllt.

Tab. 5.8: Mittlere Beanspruchung (gewichteter Gesamtwert) beim Lernen mit AC und mit dem KSM getrennt nach der Periode.

	Periode 1			Periode 2			Total (ungewichtetes Mittel)		
	M	SD	N	M	SD	N	M	SD	N
AC	42.33	10.57	12	50.45	23.17	11	46.39	17.80	23
KSM	39.36	11.24	11	40.92	19.15	12	40.14	15.54	23
Total	40.85	10.91	23	45.69	21.24	23			

Anmerkung: Die dunkelgrau eingefärbten Zahlen entstammen der Gruppe 1 (AC -> KSM), die hellgrau eingefärbten Zahlen der Gruppe 2 (KSM -> AC). Die übrigen Zahlen beziehen sich auf beide Gruppen gleichzeitig (within-subjects).

Die zweifaktorielle Varianzanalyse zeigt keine signifikante Interaktion zwischen Lernmethode und Periode ($F = 0.44$, $df = 1/42$, $p = .51$); ein Carry-Over-Effekt wird folglich nicht angenommen. Die Natur der geprüften Effekte der Varianzanalyse zeigt das Interaktionsdiagramm für den Faktor Periode in Abbildung 5.13. Betrachtet man beide Perioden separat, so ist zu sehen, dass sich die mittleren Beanspruchungswerte der beiden Lerngruppen in der ersten Periode nur wenig voneinander unterscheiden ($M_{AC} = 42.33$, $M_{KSM} = 39.36$), während in der zweiten Periode ein deutlicherer Mittelwertsunterschied vorliegt ($M_{AC} = 50.45$, $M_{KSM} = 40.92$). Aufgrund des nicht signifikanten Interaktionseffekts zwischen Lernmethode und Periode kann dieser unterschiedliche Einfluss der Lernmethode je nach Periode aber auch zufällig entstanden sein.

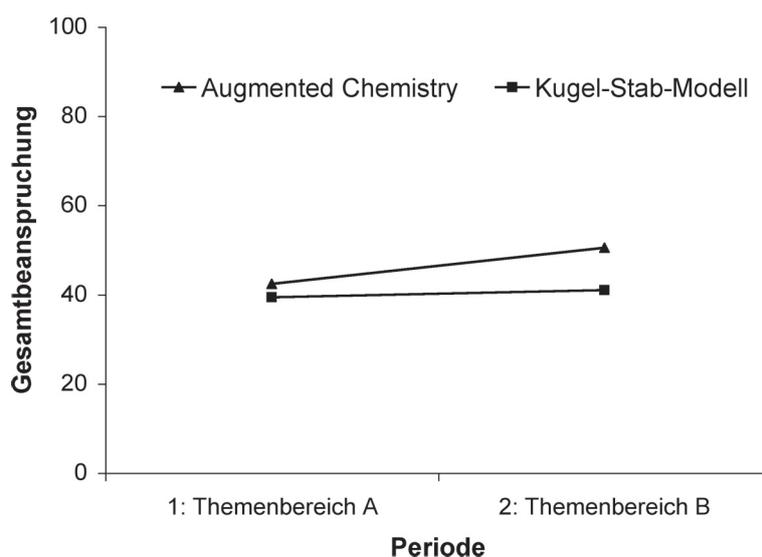


Abb. 5.13: Die mittlere Gesamtbeanspruchung pro Lernmethode und nach der Periode.

Wie in Tabelle 5.8 zu sehen ist, fühlten sich Personen - unabhängig von der untersuchten Periode-, die mit AC gelernt hatten, etwas beanspruchter ($M = 46.39$) als Personen, die mit dem KSM gelernt hatten ($M = 40.14$). Doch dieser Unterschied erweist sich im geprüften Haupteffekt der Varianzanalyse als nicht signifikant ($F = 1.59$, $df = 1/42$, $p = .22$). Der Effekt der Lernmethode ist mit $d = .37$ eher klein. Es zeigt sich somit kein relevanter Unterschied bei der empfundenen Gesamtbeanspruchung von Personen, die mit AC, und solchen, die mit dem KSM gelernt haben. Die Varianzanalyse zeigt weiter auch keinen signifikanten Haupteffekt der Periode ($F = 0.95$, $df = 1/42$, $p = .34$). Bei der Interpretation der leicht höher empfundenen Gesamtbeanspruchung in der zweiten Periode ($M = 45.69$) - unabhängig von der Lernmethode - im Vergleich zur ersten Periode ($M = 40.85$) kann der Zufall somit nicht ausgeschlossen werden.

Subskalen der Beanspruchung

Im Folgenden wird auf die mittleren Beanspruchungswerte in den sechs Subskalen des Nasa Task Load Index genauer eingegangen. Für die Auswertungen wurden die Rohwerte der Ratings in den einzelnen Subskalen verwendet ohne eine Gewichtung der einzelnen Subskalen vorzunehmen.

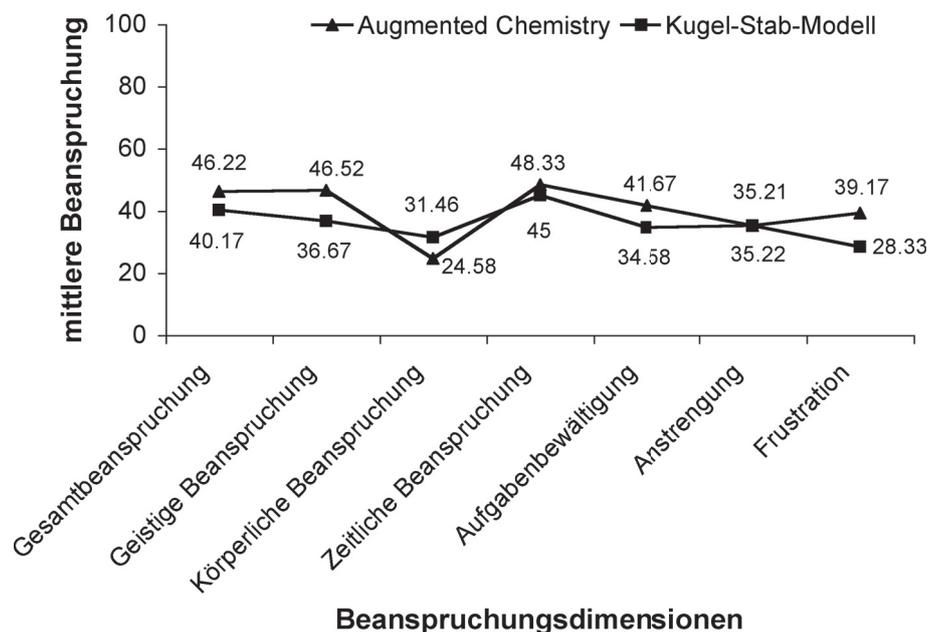


Abb. 5.14: Die durchschnittlichen Beanspruchungswerte in den sechs Subskalen des NASA-TLX getrennt nach AC und dem KSM.

Es fällt auf, dass die Beanspruchung in allen Skalen - ausser bei der körperlichen Beanspruchung - durch das Arbeiten mit AC höher war als durch das Arbeiten mit dem KSM (Abbildung 5.14). Am ausgeprägtesten zeigen sich die Unterschiede bei den Skalen *geistige Beanspru-*

chung und *Frustration* mit einer Mittelwertsdifferenz um die 10 Punkte. Allgemein liegen die durchschnittlichen Beanspruchungswerte im mittleren bis unteren Skalenbereich. Die tiefsten Werte sind unter beiden Lernmethoden bei der körperlichen Beanspruchung zu finden. Mit beiden Lernmethoden fühlten sich die Upn demnach am wenigsten durch körperliche Anstrengung beansprucht. Was auffällt ist, dass diese Dimension die einzige unter den sechs Beanspruchungsdimensionen ist, welche beim KSM im Schnitt höher bewertet wurde als bei AC. Dies ist plausibel, da mit dem KSM das Zusammenfügen und Auseinandernehmen von Kugeln und Stäbchen mehr Kraft erfordert als das einfache Klicken auf den Greifer bei AC.

Geistige Beanspruchung: Zwischen der Lernmethode und geistiger Beanspruchung zeigt sich folgender Zusammenhang: Wie in Abbildung 5.14 ersichtlich ist, fühlten sich Personen, die mit AC gelernt haben geistig beanspruchter ($M = 46.52$) als Personen, die mit dem KSM gelernt haben ($M = 36.67$). Das Lernen mit AC erforderte folglich ein höheres Mass an geistiger Arbeit und Wahrnehmungsleistungen sowie höhere Genauigkeit und wurde in dieser Hinsicht als anspruchsvoller und komplexer empfunden. Da die Varianzanalyse keinen signifikanten Interaktionseffekt zwischen Lernmethode und Periode für die geistige Beanspruchung ($F = 0.04$, $df = 1/43$, $p = .85$) zeigt, ist eine globale Interpretation der Haupteffekte zulässig. Der beschriebene Unterschied zwischen den Lernmethoden (Haupteffekt der Lernmethode) erweist sich in der Varianzanalyse als knapp signifikant ($F = 3.93$, $df = 1/43$, $p = .05$). Es handelt sich hier um einen mittleren und somit nicht unbedeutenden Effekt der Lernmethode von $d = .57$. Somit kann die Aussage, dass das Lernen mit AC zu einer höheren geistigen Beanspruchung führt als das Lernen mit dem KSM, verallgemeinert werden, und der Effekt ist praktisch bedeutsam.

Frustration: Allgemein war das Frustrationsniveau bei den Lernenden eher gering. Es zeigt sich aber, dass Personen, die mit AC gelernt haben, über ein höheres Frustrationsniveau verfügten ($M = 39.17$) als Personen, die mit dem KSM gelernt haben ($M = 28.33$). Durch den nicht signifikanten Interaktionseffekt der Varianzanalyse können die Haupteffekte global interpretiert werden ($F = 0.23$, $df = 1/44$, $p = .63$). Der beschriebene Unterschied zwischen den Lernmethoden in der empfundenen Frustration erweist sich in der Varianzanalyse als nicht signifikant ($F = 1.92$, $df = 1/44$, $p = .17$). Der Effekt ist mit $d = .40$ eher schwach, woraus geschlossen werden kann, dass die Lernmethode einen eher kleinen, aber nicht generalisierbaren Zusammenhang mit dem Frustrationsempfinden hat.

Es wurden hier lediglich die zwei Dimensionen *geistige Beanspruchung* und *Frustration* detaillierter betrachtet, da sie die ausgeprägtesten Unterschiede aufzeigen. In Tabelle 5.9 ist ersichtlich, dass ausser der geistigen Beanspruchung keine der sechs Dimensionen einen signifikanten Haupteffekt der Lernmethode aufweist. Die sechs Dimensionen sind in Tabelle 5.9 in eine Rangreihe gebracht und zwar von der Dimension mit der höchsten zur Dimension mit der geringsten mittleren Beanspruchungsbewertung. Die jeweiligen Mittelwerte für AC und das

KSM sind aus der Abbildung 5.14 zu entnehmen. Wie in Tabelle 5.9 ersichtlich ist, wurde die zeitliche Beanspruchung am höchsten beurteilt. Das Gefühl, unter Zeitdruck beim Lösen der Aufgaben zu stehen, war unter den Dimensionen der belastendste Aspekt. Weit geringer wurde die körperliche Beanspruchung eingeschätzt. Am wenigsten fühlten sich die Upn also durch körperliche Arbeit wie Drücken, Ziehen, Drehen oder Steuern beansprucht.

Tab. 5.9: Beanspruchung der Lernenden in den sechs Subdimensionen mit ANOVA-Analyse auf Unterschiede zwischen den Lernmethoden (Prüfung auf einen Behandlungseffekt).

Dimensionen	M	SD	N	F	p
Zeitliche Beanspruchung	46.67	24.61	48	0.22	.64
Geistige Beanspruchung	41.49	17.26	47	3.93	.05
Aufgabenbewältigung	38.13	27.01	48	0.84	.37
Anstrengung	35.21	20.67	47	0.00	.99
Frustration	33.75	27.08	48	1.92	.17
Körperliche Beanspruchung	28.02	22.19	48	1.23	.27

Anmerkungen: Skala von 0 bis 100. Unterschiede in N aufgrund fehlender Angaben.

5.4 Emotionale Stimmung

In den folgenden Kapiteln wird die emotionale Stimmung der Upn unmittelbar nach dem Arbeiten mit dem jeweiligen Lernsystem unter den drei Aspekten *Aktiviertheit*, *Valenz* und *Dominanz* genauer untersucht. Es geht darum, herauszufinden, ob sich Personen, die mit AC gearbeitet haben, von Personen, die mit dem KSM gearbeitet haben, hinsichtlich dieser drei emotionalen Dimensionen unterscheiden. Vorerst werden wieder als Übersicht die Häufigkeitsverteilungen der Upn auf die drei Dimensionen der emotionalen Stimmung graphisch dargestellt.

5.4.1 Verteilung der Untersuchungspersonen hinsichtlich ihrer emotionalen Stimmung

Aktiviertheit

In Abbildung 5.15 ist die Verteilung der Aktiviertheit der Upn nach den Lernphasen über beide Perioden hinweg abgebildet. Die Schiefe dieser Verteilung beträgt 0.05 und die Steilheit -0.29. Da beide Werte annähernd gleich Null sind, kann von einer Normalverteilung ausgegangen werden (Tabachnik & Fidell, 1996). Durchschnittlich fühlten sich die Upn nach dem Lernen mässig aktiviert. Sie waren eher ruhig und wenig emotional erregt, was sich im Mittelwert von 4.38 zeigt, welcher sich um die Skalenmitte befindet (Wertebereich der Skala: von 1 bis 9).

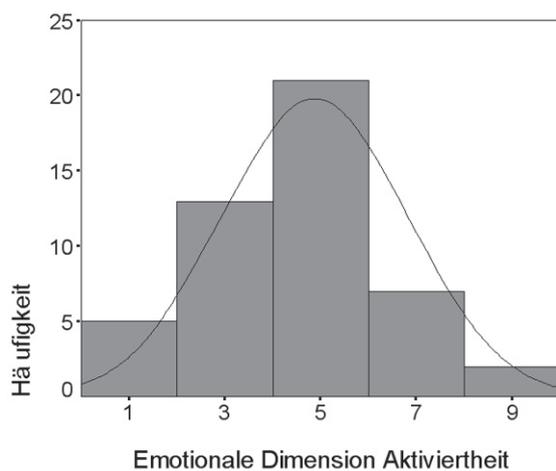


Abb. 5.15: Häufigkeitsverteilung der Aktiviertheit bei den Upn über beide Perioden hinweg ($M = 4.38$, $SD = 1.93$, $N = 48$).

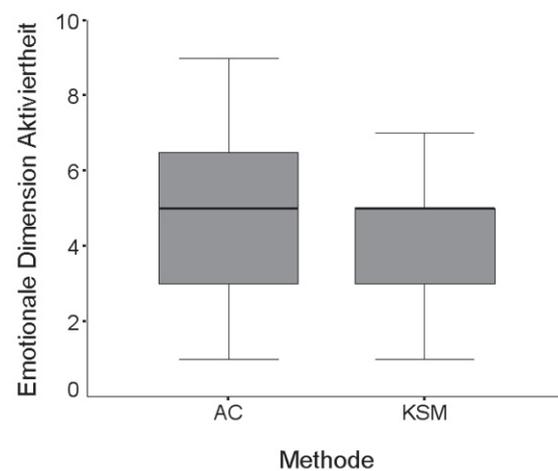


Abb. 5.16: Verteilung der Aktiviertheit getrennt nach AC und dem KSM ($N_{AC} = N_{KSM} = 24$).

Wie die beiden Verteilungen der Aktiviertheit getrennt nach AC und dem KSM in Abbildung 5.16 zeigen, so unterscheiden sich die emotionalen Reaktionen auf die beiden Lernmethoden im Schnitt nicht, wie aus dem identischen Median von $M_d = 5$ ersichtlich ist. Im Schnitt fühlten sich die Upn daher, ob sie mit AC oder dem KSM gelernt haben, mässig emotional erregt. Die Werte streuen beim Lernen mit AC aber stärker als beim Lernen mit dem KSM, insbesondere im oberen Bereich. Während bei AC 50% aller Werte zwischen 3 und 6.75 Punkten liegen, so finden sich beim KSM 50% aller Werte zwischen 3 und 5 Punkten. 75% aller Werte liegen beim KSM unterhalb von 5 Punkten, bei AC hingegen sind unterhalb dieses Wertepunkts nur 50% aller Werte zu finden. Bei AC reagierten folglich die Upn mit mehr erregungsbezogenen Gefühlen als beim KSM.

Valenz

Bei der Verteilung der Valenz der Emotion bei den Upn über beide Perioden hinweg (Abbildung 5.17) zeigt sich, dass die Lernsituation bei den Upn ein eher positives Gefühl ausgelöst hat. Die Upn waren im Schnitt nach dem Lernen von einem angenehmen Gefühlszustand betroffen, was sich durch den Mittelwert von 6.85 zeigt. Mit einer Schiefe von -0.71 und einer Steilheit von 0.40 handelt es sich um eine leicht linksschiefe Verteilung. Die Werte liegen aber beide annähernd bei Null, wodurch bei den folgenden Auswertungen von einer annähernden Normalverteilung ausgegangen wird. Mit einer quadratischen Transformation konnte die Verteilung zwar einer Normalverteilung stärker angenähert werden, wie sich dann aber zeigte, würden sich die Resultate der Varianzanalyse der transformierten und der nicht-transformierten Daten kaum voneinander unterscheiden, weshalb von einer Datentransformation abgesehen wird.

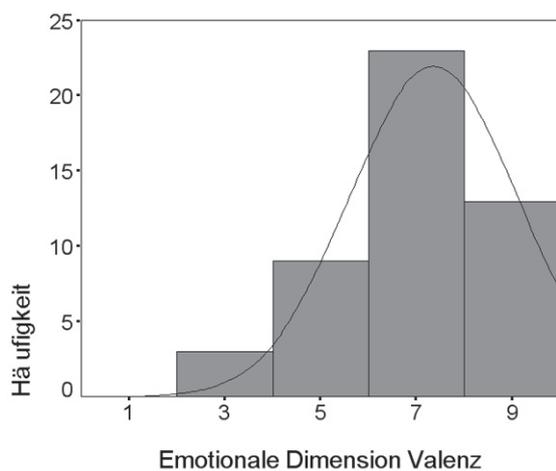


Abb. 5.17: Häufigkeitsverteilung der Valenz bei den Upn über beide Perioden hinweg ($M = 6.85$, $SD = 1.74$, $N = 48$).

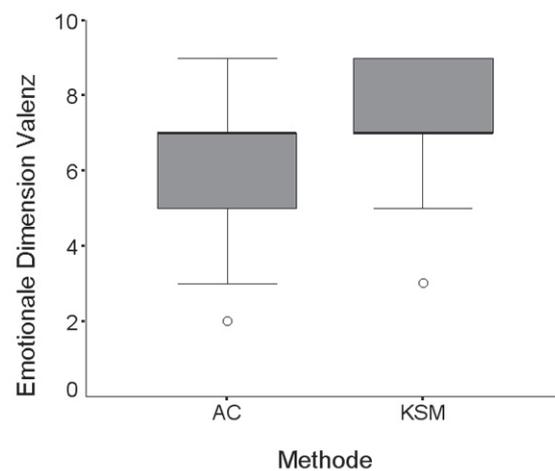


Abb. 5.18: Verteilung der Valenz getrennt nach AC und dem KSM ($N_{AC} = N_{KSM} = 24$)

Betrachtet man die Verteilungen getrennt nach AC und dem KSM (Abbildung 5.18), so zeigt sich kein Unterschied in den mittleren Valenzurteilen. Bei beiden Lernmethoden findet sich ein identischer Median von $Md = 7$. Dieser Wert im oberen Skalenbereich bedeutet, dass die Lernsituation im Schnitt bei den Upn ein positives Gefühl ausgelöst hat. Beim Lernen mit AC finden sich aber mehr Werte im unteren Bereich, während sich beim Lernen mit dem KSM im oberen Skalenbereich die meisten Werte finden. Die Antworten bei AC variieren zwischen 3 und 9 Punkten, beim KSM zwischen 5 und 9 Punkten, mit Ausreißern nach unten bei beiden Lernmethoden. Beim Lernen mit AC liegen 50% aller Werte zwischen 5 und 7 Punkten, beim Lernen mit dem KSM hingegen findet man 50% aller Werte zwischen 7 und 9 Punkten. Offenbar löste das Lernen mit AC im Vergleich zum Lernen mit dem KSM auch negativere, unangenehmere Gefühle aus, während auf das Lernen mit dem KSM mehr Personen als bei AC auch mit sehr positiven, angenehmen Gefühle reagierten.

Dominanz

Die Häufigkeitsverteilung des Dominanzempfindens über beide Perioden hinweg fällt leicht linksschief aus (Abbildung 5.19). Die Schiefe beträgt -0.42 und die Steilheit -0.25 . Beide Werte sind annähernd gleich Null, wodurch von einer annähernden Normalverteilung ausgegangen wird. Die Situation nach dem Lernen wurde im Schnitt relativ gut unter Kontrolle empfunden, was der Mittelwert von 6.10 zeigt. Die ganze Skalenbreite wird ausgenutzt.

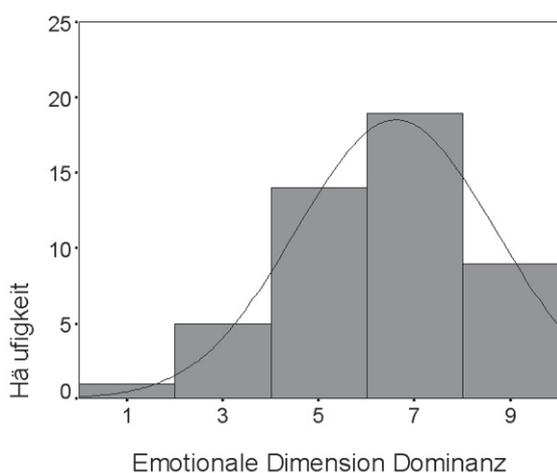


Abb. 5.19: Häufigkeitsverteilung der Dominanz bei den Upn über beide Perioden hinweg ($M = 6.10$, $SD = 2.07$, $N = 48$).

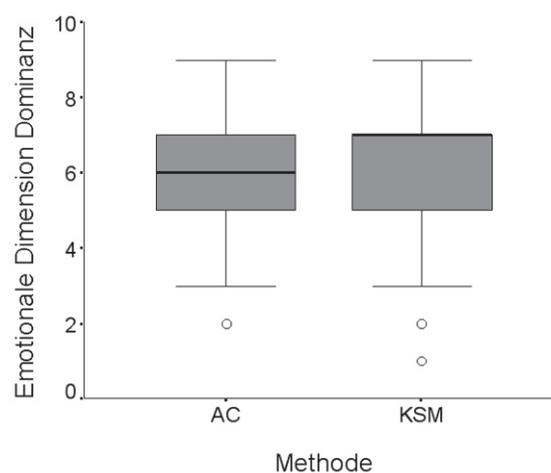


Abb. 5.20: Verteilung der Dominanz getrennt nach AC und KSM ($N_{AC} = N_{KSM} = 24$)

Die Verteilungen getrennt nach AC und dem KSM (Abbildung 5.20) zeigen, dass nach dem Lernen mit dem KSM die Situation im Schnitt etwas besser unter Kontrolle empfunden wurde als nach dem Lernen mit AC ($Md_{KSM} = 7$; $Md_{AC} = 6$). Die Antworten der Upn variieren alle zwischen 5 und 9 Punkten, mit einigen Ausreißern nach unten. Während beim KSM 50% aller Daten oberhalb vom Wert 7 liegen, so finden sich bei AC nur 25% aller Daten oberhalb dieses Wertes. Ansonsten gibt es zwischen den beiden Lernmethode keine wesentlichen Unterschiede.

5.4.2 Unterschiede in der emotionalen Stimmung bei Personen, die mit Augmented Chemistry, und bei Personen, die mit dem Kugel-Stab-Modell gelernt haben

Tabelle 5.10 gibt die Mittelwerte und Standardabweichungen der Aktiviertheit, Valenz und Dominanz für Personen, die mit AC, und Personen, die mit dem KSM gearbeitet haben und separat für beide Perioden wieder.

Tab. 5.10: Die Mittelwerte in den drei Dimensionen der emotionalen Stimmung für Personen, die mit AC und jene, die mit dem KSM gearbeitet haben getrennt nach der Periode.

		Periode 1			Periode 2			Total (ungewichtetes Mittel)		
		M	SD	N	M	SD	N	M	SD	N
Aktiviertheit	AC	4.17	2.17	12	5.25	2.09	12	4.71	2.16	24
	KSM	4.58	1.44	12	3.50	1.73	12	4.04	1.65	24
	Total	4.38	1.81	24	4.38	2.08	24			
Valenz	AC	7.00	1.21	12	6.08	2.31	12	6.54	1.87	24
	KSM	6.92	1.62	12	7.42	1.56	12	7.17	1.58	24
	Total	6.96	1.40	24	6.75	2.05	24			
Dominanz	AC	6.25	1.60	12	5.83	2.52	12	6.04	2.07	24
	KSM	6.08	1.78	12	6.25	2.45	12	6.17	2.10	24
	Total	6.17	1.66	24	6.04	2.44	24			

Anmerkung: dunkelgrau eingefärbte Zahlen entstammen der Gruppe 1 (AC -> KSM), hellgrau eingefärbten Zahlen der Gruppe 2 (KSM -> AC), die übrigen Zahlen beziehen sich auf beide Gruppen.

Aktiviertheit

Aus Tabelle 5.10 und Abbildung 5.21 ist ersichtlich, dass in Periode 1 die durchschnittliche Aktiviertheit nach dem Lernen mit dem KSM leicht höher war ($M = 4.58$) als nach dem Lernen mit AC ($M = 4.17$), während es sich in Periode 2 genau umgekehrt verhält: Bei Personen, die mit dem KSM gelernt hatten, fiel die Aktiviertheit durchschnittlich deutlich geringer aus ($M = 3.50$) als bei Personen, die mit AC gelernt hatten ($M = 5.25$). Dies zeugt von einem Interaktionseffekt, welcher bei der zweifaktoriellen Varianzanalyse auch signifikant ausfällt ($F = 3.98$, $df = 1/44$, $p = .05$). Man kann folglich nicht generell aussagen, welches Lernsystem zu einer höheren Aktiviertheit führt, es kann nur eine differenzierte Betrachtung wie oben beschrieben erfolgen.

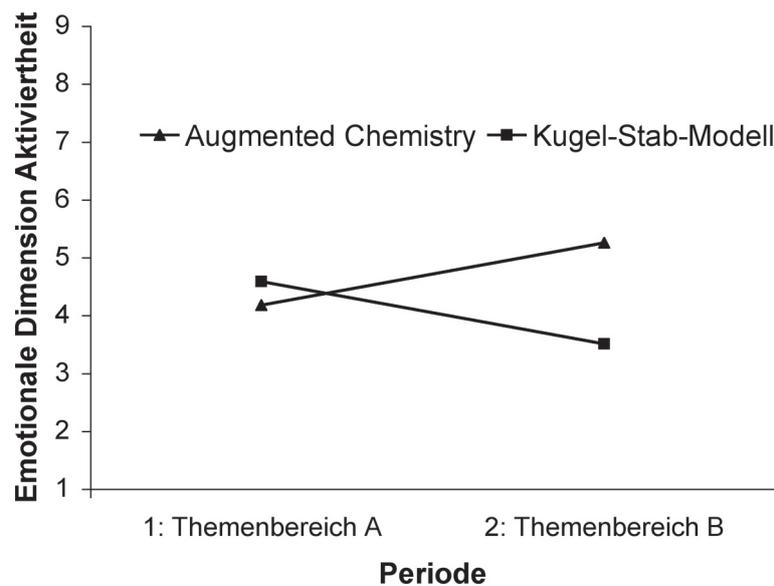


Abb. 5.21: Die mittlere Aktiviertheit pro Lernmethode und nach der Periode.

Der Interaktionseffekt kann als ein psychologischer Carry-Over-Effekt interpretiert werden und äussert sich dabei folgendermassen: Mit dem KSM wurde in der zweiten Periode (Gruppe 1) im Mittel mit deutlich weniger erregungsbezogenen Emotionen reagiert, wenn zuvor mit AC gearbeitet wurde als mit dem KSM ohne vorherige Lernphase (Gruppe 2). Das Arbeiten mit AC führte hingegen in der zweiten Periode zu einem deutlich höheren durchschnittlichen Aktiviertheitsniveau, wenn zuvor mit dem KSM gearbeitet wurde als das Arbeiten mit AC ohne vorherige Lernphase. Der Mittelwertsunterschied in der zweiten Periode ist ausserdem deutlich stärker ausgeprägt. Handelt es sich hier um einen Carry-Over-Effekt, so würde dies bedeuten, dass das Arbeiten mit AC eine Senkung des Aktiviertheitsniveaus in der folgenden KSM-Lerneinheit bewirkt, während das Arbeiten mit dem KSM eher eine Steigerung der Aktiviertheit in der folgenden AC-Lerneinheit hervorruft. Ferner ist zu erkennen, dass das durchschnittliche Aktiviertheitsniveau nach dem Lernen mit AC bei beiden Gruppen höher lag als nach dem Lernen mit dem KSM. Es könnte daher auch sein, dass man es hier mit einem Gruppeneffekt zu tun hat. Worauf der Interaktionseffekt zurückgeführt werden kann, ist keine einfach zu klärende Frage und wird in Kapitel 6 genauer zu erläutern versucht. Der signifikante Interaktionseffekt verunmöglicht eine globale Interpretation der beiden Haupteffekte, sofern sie signifikant ausfallen. Es zeigt sich aber, dass weder der Behandlungseffekt ($F = 1.51$, $df = 1/44$, $p = .23$) noch der Periodeneffekt ($F = 0.00$, $df = 1/44$, $p = 1.00$) signifikant sind. Die Aktiviertheitswerte nach dem Arbeiten mit AC unterscheiden sich also nicht signifikant von den nach dem Arbeiten mit dem KSM ermittelten Werten. Die Lernmethode steht folglich mit der Aktiviertheit in keinem Zusammenhang, der nicht auch durch den Zufall erklärt werden könnte.

Valenz

Während sich in der ersten Periode die beiden Lerngruppen kaum in ihren Valenzurteilen voneinander unterscheiden ($M_{AC} = 7.00$, $M_{KSM} = 6.92$), so ist dieser Mittelwertsunterschied in der zweiten Periode deutlich stärker zu Gunsten des KSM ausgeprägt ($M_{AC} = 6.08$, $M_{KSM} = 7.42$) (Tabelle 5.10). Die zweifaktorielle Varianzanalyse für die angegebenen Valenzen ergibt keinen signifikanten Interaktionseffekt ($F = 2.03$, $df = 1/44$ $p = .16$).

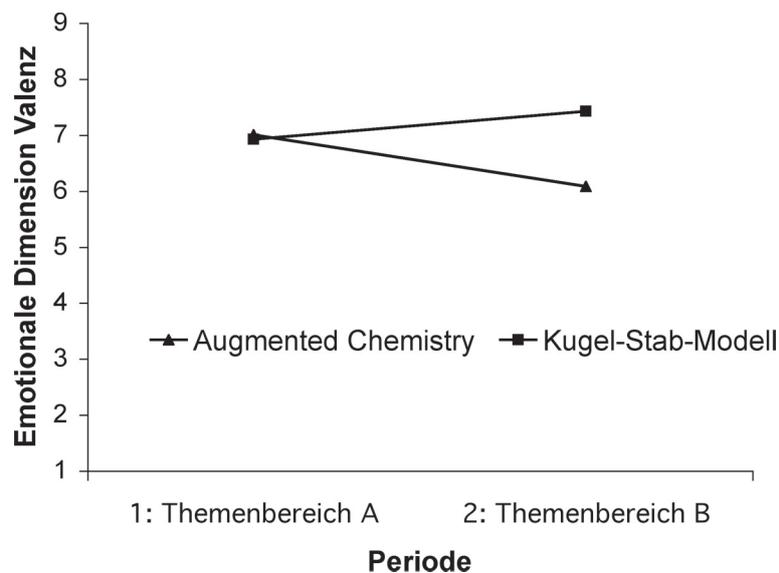


Abb. 5.22: Die mittlere Valenz pro Lernmethode und nach der Periode.

Wie in Tabelle 5.10 und Abbildung 5.22 ersichtlich ist, wurde bei Personen, welche mit dem KSM gelernt hatten - unabhängig von der untersuchten Periode - im Schnitt ein positiveres Gefühl ($M = 7.17$) ausgelöst als bei Personen, welche mit AC gelernt hatten ($M = 6.54$). Der Haupteffekt der Lernmethode erweist sich in der Varianzanalyse aber als nicht signifikant ($F = 1.58$, $df = 1/44$, $p = .22$). Die Valenzwerte nach dem Arbeiten mit AC unterscheiden sich demnach nicht signifikant von den nach dem Arbeiten mit dem KSM ermittelten Werten. Der Effekt ist mit $d = .36$ schwach, woraus geschlossen werden kann, dass die Lernmethode einen kleinen, aber nicht generalisierbaren Zusammenhang mit der Valenz hat.

Des Weiteren zeigt sich, dass in der ersten Periode - unabhängig von der Lernmethode - die Valenz etwas höher war ($M = 6.96$) als in der zweiten Periode ($M = 6.75$), aber nur in geringem Ausmass. Der Haupteffekt der Periode ist auch nicht signifikant ($F = 0.18$, $df = 1/44$, $p = .68$) und der Effekt ist mit $d = .12$ praktisch irrelevant. Folglich steht sowohl die Lernmethode als auch die Periode mit der empfundenen Valenz in keinem Zusammenhang, der nicht auch durch den Zufall erklärt werden könnte. Die Voraussetzung homogener Stichprobenvarianzen für die Varianzanalyse wurde für die Variable Valenz nicht erfüllt.

Dominanz

Die zweifaktorielle Varianzanalyse für die emotionale Dimension Dominanz zeigt, dass keiner der geprüften Effekte signifikant ist (Interaktionseffekt: $F = 0.23$, $df = 1/44$, $p = .64$; Behandlungseffekt: $F = 0.04$, $df = 1/44$, $p = .84$; Periodeneffekt: $F = 0.04$, $df = 1/44$, $p = .84$). Der Effekt für den Unterschied in den Lernmethoden ist mit $d = .06$ praktisch irrelevant. Es konnte somit kein relevanter Unterschied im Dominanzempfinden zwischen Personen, die mit AC, und solchen, die mit dem KSM gelernt haben, gefunden werden. Die Mittelwerte befinden sich alle im oberen Skalenbereich (Abbildung 5.23). Im Schnitt hatten die Upn also das Gefühl, die Situation nach der Lernphase relativ gut unter Kontrolle zu haben. Unterschiede zwischen den Lernmethoden zeigen sich kaum.

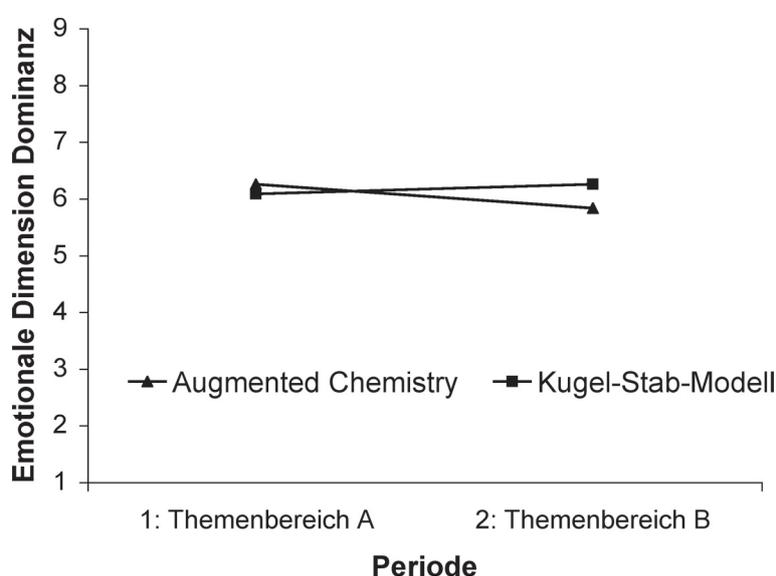


Abb. 5.23: Das mittlere Dominanzempfinden pro Lernmethode und nach der Periode.

5.5 Zufriedenheit beim Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell

Im Folgenden wird auf den Fragebogen “Lernen mit Augmented Chemistry (oder Kugel-Stab-Modell)” genauer eingegangen. Dabei geht es um eine Einschätzung der Zufriedenheit der Lernenden beim Arbeiten mit den beiden Lernsystemen. In Anbetracht der Vielzahl an erhobenen Daten wird an dieser Stelle auf die Auswertung des Semantischen Differentials verzichtet. Analysiert werden aber die 12 Item-Aussagen (Kapitel 5.5.1), die offenen Fragen (Kapitel 5.5.2) und die Beurteilung der Lernsysteme auf einer Notenskala (Kapitel 5.5.3).

5.5.1 Unterschiede in der Zufriedenheit beim Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell

Die Resultate zu den 12 Item-Aussagen aus dem Fragebogen zur Zufriedenheit sind in Tabelle 5.11 ersichtlich.

Tab. 5.11: Zufriedenheit der Upn beim Lernen mit AC bzw. dem KSM mit ANOVA-Analyse auf Unterschiede zwischen den Lernmethoden.

Items	M _{AC}	SD _{AC}	M _{KSM}	SD _{KSM}	F	p
1 Das Arbeiten mit AC/KSM hat Spass gemacht	3.67	1.24	4.04	.69	1.70	.20
2 Mit Hilfe von AC/KSM Fachkenntnisse erweitern können	2.96	1.20	3.00	.98	.02	.89
3 Gleichen Lerninhalt auch ohne AC/KSM erarbeiten	2.25	.85	2.42	1.10	.34	.56
4 AC/KSM hat geholfen Lerninhalt besser zu verstehen	2.79	1.02	3.38	.71	5.53	.02
5 AC/KSM war gute Unterstützung beim Lösen der Aufgaben	3.25	1.26	3.83	1.01	3.18	.08
6 Mit AC/KSM macht Lernen mehr Spass	3.71	1.23	4.08	.78	1.53	.22
7 Durch AC/KSM Interesse am Lerninhalt gesteigert	3.17	1.13	3.42	.78	.80	.38
8 AC/KSM auch in Zukunft gerne anwenden	3.50	1.25	3.75	.94	.62	.44
9 AC/KSM auch anderen Studierenden empfehlen	3.54	1.32	3.58	.97	.02	.90
10 AC/KSM ist nützliches Instrument zum Lernen	3.50	1.35	3.71	.81	.41	.52
11 Auf Lerninhalte konzentrieren	2.88	.99	3.42	1.10	3.08	.09
12 Gut alleine mit AC/KSM zurechtkommen	3.67	1.31	4.08	.93	1.72	.20

Anmerkungen: 5-stufige Skala von 1 ("stimme gar nicht zu") bis 5 ("stimme völlig zu"), Item 3 wurde umgepolt (N = 24; df = 1/44)

Auch wenn sich die Antworten zwischen den beiden Lernsystemen im Mittel kaum voneinander unterscheiden, so zeigt sich durchwegs eine leicht bessere Bewertung der Items beim KSM. Beim Lernen mit dem KSM waren die Upn folglich leicht zufriedener als beim Lernen mit AC. Ein signifikanter Unterschied in den Item-Antworten findet sich aber nur bei Item 4: Ob das Lernsystem geholfen hat den Lerninhalt besser zu verstehen, dem stimmten durchschnittlich viel mehr die Lernenden mit dem KSM (M = 3.38) als die Lernenden mit AC zu (M = 2.79) (F = 5.53, df = 1/44, p < .05). Bei den übrigen Items sind die Unterschiede zu Gunsten des KSM nur gering.

Interessant sind die Antworten zu den Items 1 und 6, welche sich inhaltlich beide auf den Spassfaktor der Lernsituation beziehen. Dieser Aspekt von Spass beim Lernen taucht in verschiedenen Fragen der Untersuchung auf, so auch in der Präferenzbefragung (Kapitel 5.6) aber auch implizit in der Valenzbeurteilung bei den Emotionen (Kapitel 5.4.2). Im Fragebogen zur Zufriedenheit wurden die zwei Items zu Spass bei beiden Lernmethoden am höchsten von allen Items bewertet. Die Lernenden mit dem KSM sind sich hier aber einiger als die Lernenden

mit AC, was aus der grösseren Standardabweichung bei AC hervorgeht. Was die Effektivität des Lernens angeht (Items 2, 3, 4), so wurden diese Items am schlechtesten bewertet. Bei allen Items fanden sich aber praktisch alle Itemantworten über dem arithmetischen Mittel von 2.5. Im Schnitt waren die Upn somit zufrieden bis sehr zufrieden mit den beiden Lernmethoden.

5.5.2 Berichtete positive und negative Aspekte zum Lernen mit Augmented Chemistry und mit dem Kugel-Stab-Modell

Im Fragebogen "Das Lernen mit Augmented Chemistry (oder Kugel-Stab-Modell)" hatten die Upn die Möglichkeit, zu notieren, was ihnen beim Arbeiten mit den beiden Lernsystemen positiv und was ihnen negativ aufgefallen war. Die Antworten lassen sich in folgende Bereiche einordnen:

- a) Aspekte zur Bedienung und Handhabung des Lernsystems
- b) Aspekte zum Lernsystem an sich
- c) Aspekte zur Darstellung der Moleküle
- d) allgemeine affektive Eindrücke

Zu a): Während es beim KSM eher als mühsam empfunden wurde, die Kugeln und Stäbchen zu einem Molekülmodell zusammenzufügen, so wurde das Konstruieren von Molekülen bei AC als relativ einfach und schnell beurteilt. Insbesondere das Bauen grosser Moleküle bereitete mit dem KSM Mühe. Vielleicht gerade deshalb, weil Teile oftmals wieder auseinanderfielen. Dies kann bei AC nicht passieren, was einen Vorteil für das computergestützte Lernsystem ist. Nun gibt es aber auch bei AC einige Probleme bei der Bedienung. Hauptsächlich lagen diese gemäss den Upn bei der Aktivierung von Doppelbindungen, welche über die richtige Positionierung des Würfels vorgenommen werden muss. Dies war oft mit Feinarbeit verbunden, da die gewünschte Bindungsart (Stelle auf dem Tetraeder) teilweise schwer zu fassen war (anstatt sich auf der Kante des Tetraeders zu positionieren, hüpfte die anvisierte Bindungsstelle zwischen den Ecken hin und her.).

Zu b): Bei beiden Lernsystemen wurde das praktische Arbeiten sehr geschätzt. Bei AC war das Arbeiten am Computer ein häufig genannter Pluspunkt und auch die Technologie faszinierte viele. Die verschiedenen Möglichkeiten, die AC bietet, wurden als sehr positiv wahrgenommen. Beim KSM wurde das Material teilweise als veraltet empfunden. Nun gibt es bei AC noch eine Reihe technischer Probleme. Das System reagierte teilweise nicht schnell genug oder sogar falsch auf Eingaben. So wurden die ausgewählten Atome aus dem Buch vom System teilweise verwechselt: Anstatt Wasserstoff (H) wurde manchmal ein Stickstoff-Atom (N) erfasst. Auch die Modus-Karten wurden gelegentlich nicht sofort erkannt. Die Funktionsschritte mussten dann wiederholt werden bis das System die Aktion erkannte und richtig reagierte. Was die

gesprochenen Lerntexte betrifft, so konnten diese zwar beliebig wiederholt abgespielt werden, aber jeweils nur wieder von Anfang an. Ein gezieltes Zurückspulen wäre sinnvoller und wurde von den Upn auch gewünscht. Das Arbeiten mit dem Greifer wurde teilweise als mühsam empfunden; er war vielen zu umständlich und “klobig”. Auch das Design von AC wurde kritisiert. So wünschten sich einige eine “jüngere” und “frischere” Aufmachung.

Zu c): Die Antworten unterscheiden sich kaum zwischen den beiden Lernmethoden. Die dreidimensionale Veranschaulichung der Moleküle wurde bei beiden Lernsystemen als sehr gut empfunden. Die Vorstellungskraft wurde gemäss den Upn positiv unterstützt, so dass sie sich die teilweise komplexen Moleküle gut vorstellen konnten.

Zu d): Bei beiden Lernsystemen wurde geschätzt, dass etwas Neues ausprobiert und zur Abwechslung eine andere Art von Unterricht durchlaufen werden konnte. Auch der Spassfaktor spielte bei beiden Lernsystemen eine wichtige Rolle. Bei den negativen Punkten waren die Antworten für AC und das KSM teilweise ähnlich. So wurde das Arbeiten mit beiden Lernmethoden als sehr zeitaufwändig und gewöhnungsbedürftig empfunden. Das Arbeiten mit AC wurde manchmal eher als ein Spielen aufgefasst. Auch beim KSM wurde der effektive Nutzen für das Lernen nicht erkannt, das Arbeiten wurde sogar als ineffizient bezeichnet. Einen Lerneffekt durch die Lernmethoden wurde von den Upn folglich nicht wirklich gesehen.

5.5.3 Benotung von Augmented Chemistry und dem Kugel-Stab-Modell

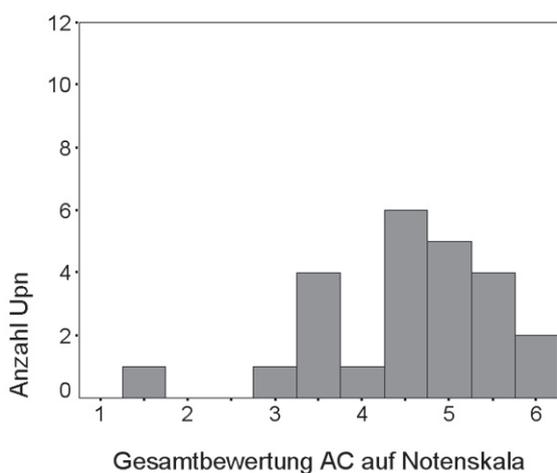


Abb. 5.24: Häufigkeitsverteilung der Benotung von AC über beide Perioden hinweg ($M = 4.5$, $SD = 1.04$, $N = 24$).

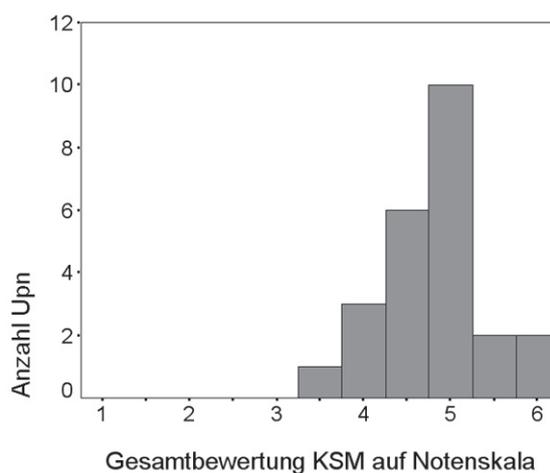


Abb. 5.25: Häufigkeitsverteilung der Benotung vom KSM über beide Perioden hinweg ($M = 4.8$, $SD = 0.59$, $N = 24$).

In Abbildung 5.24 und 5.25 sind die beiden Verteilungen der Bewertung von AC bzw. vom KSM auf einer Notenskala von 1 ('ungenügend') bis 6 ('sehr gut') dargestellt. Die Bewertung der beiden Lernsysteme auf der Notenskala fiel leicht zu Gunsten des KSM aus. Im Schnitt erhielt AC die Note 4.5 (Abbildung 5.24), das KSM erzielte durchschnittlich die leicht bessere Note von 4.8 (Abbildung 5.25). Bei AC zeigt sich aber eine deutlich stärkere Streuung der Verteilung als beim KSM. Die Lernenden mit dem KSM sind sich offenbar einiger in dieser Beurteilung als jene mit AC.

Die zweifaktorielle Varianzanalyse zeigt, dass der Interaktionseffekt der beiden Faktoren Lernmethode und Periode nicht signifikant ausfällt ($F = 0.77$, $df = 1/44$, $p = .39$). Die Natur der Effekte ist dem Interaktionsdiagramm für den Faktor Periode in Abbildung 5.26 zu entnehmen.

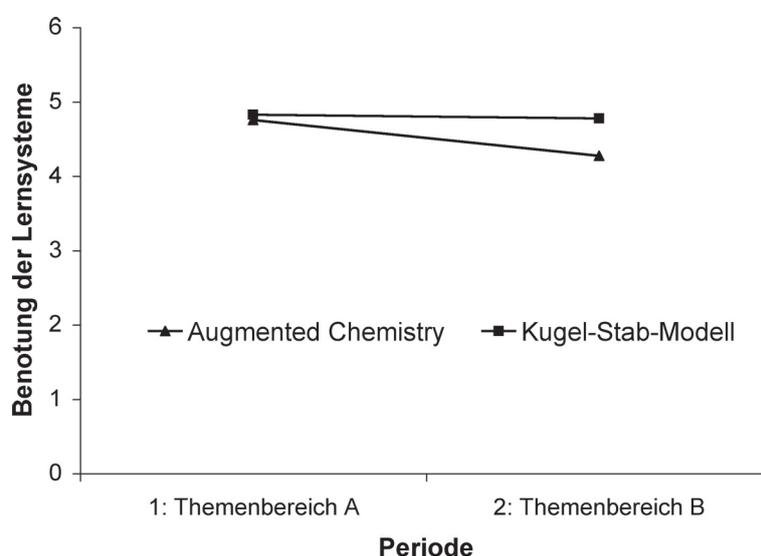


Abb. 5.26: Die durchschnittliche Bewertung der Lernmethoden auf einer Notenskala pro Lernmethode und nach der Periode.

Auch hier zeigt sich wieder, dass der Mittelwertsunterschied in der zweiten Periode stärker ist als in der ersten Periode. Insgesamt sind die Unterschiede in der Bewertung der beiden Lernmethoden aber sehr gering. Die Varianzanalyse zeigt, dass weder ein signifikanter Haupteffekt der Lernmethode ($F = 1.36$, $df = 1/44$, $p = .25$) noch ein signifikanter Haupteffekt der Periode ($F = 1.13$, $df = 1/44$, $p = .29$) vorliegt. Die Notenwerte von Lernenden mit AC unterscheiden sich demnach nicht signifikant von den nach dem Lernen mit dem KSM ermittelten Werten. Der Effekt der Lernmethode ist mit $d = .35$ schwach. Es zeigt sich also kein relevanter Unterschied bei der subjektiven Bewertung der beiden Lernsysteme auf einer Notenskala.

5.6 Präferenz für ein Lernsystem

In Abbildung 5.27 sind die neun Kriterien aus dem Fragebogen zur Präferenz eines Lernsystems abgebildet, anhand derer die Upn ihr präferiertes Lernsystem angeben mussten. In der Abbildung sind die relativen Häufigkeiten, mit denen AC oder das KSM als präferierte Lernmethode angegeben wurde in Prozent dargestellt. “Unentschieden” steht für jene Personen, welche sich für keines der beiden Lernsysteme entscheiden konnten.

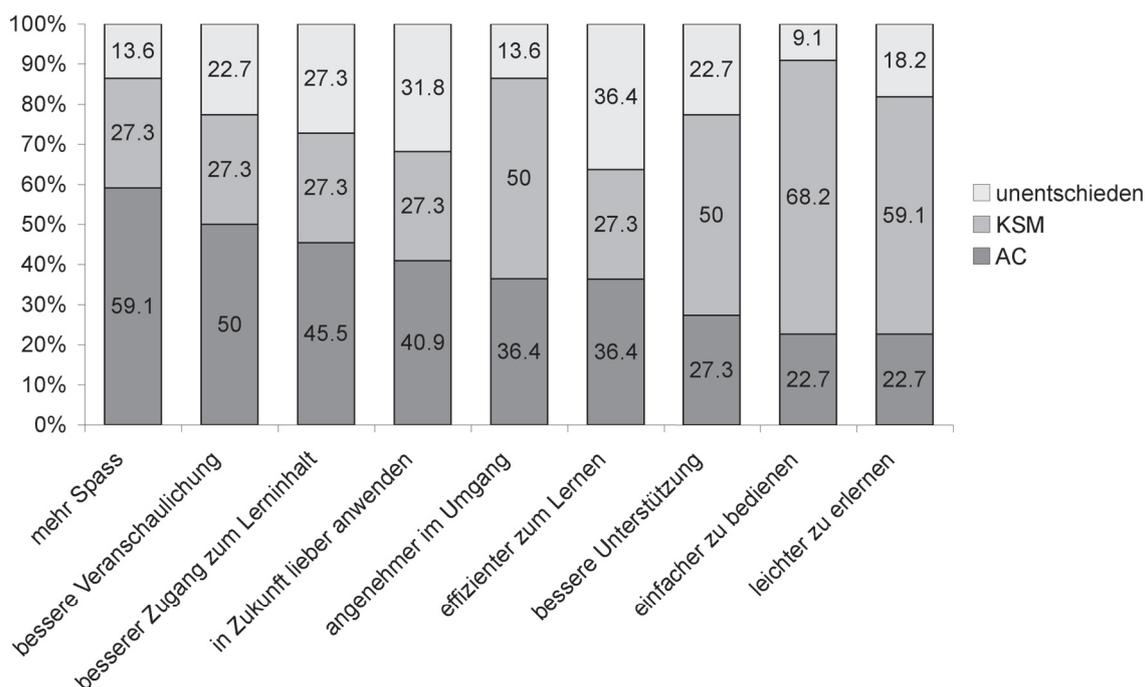


Abb. 5.27: Relative Häufigkeiten der Präferenzangaben zu den Aspekten aus dem Fragebogen zur Präferenz eines Lernsystems.

Es zeigt sich, dass die Anzahl unentschiedener Personen häufig einen beachtlichen Anteil ausmacht. In vielen Bereichen, insbesondere bei der Bewertung der Effizienz und der Verdeutlichung der Absicht einer zukünftigen Anwendung, konnten sich die Upn nicht festlegen und keine klare Präferenz für ein Lernsystem abgeben. Am eindeutigsten fällt das Bild bei der Beurteilung der Einfachheit der Bedienung aus. 68.2% der Upn fanden das KSM einfacher zu bedienen als AC, im Gegensatz zu nur 22.7%, welche AC in der Bedienung einfacher fanden. Ähnlich fällt es bei der Erlernbarkeit aus: Die Mehrheit fand das KSM leichter zu erlernen (59.1%) und angenehmer im Umgang (50%), auch bei der Unterstützungsfunktion bei den Lernaufgaben gaben die Upn bevorzugt das KSM an (50%). Es zeigt sich somit, dass zu Aspekten der Bedienbarkeit und Handhabung eine klare Präferenz für das KSM vorliegt. Anders ist es bei Aspekten wie dem Spassfaktor und der Veranschaulichung. 59.1% der Upn machte

das Lernen mit AC mehr Spass als das Lernen mit dem KSM. Bei der Veranschaulichung der Lerninhalte zeigt sich ebenfalls eine Präferenz für AC: 50% waren der Meinung, AC könne den Lerninhalt besser veranschaulichen und verständlicher machen als das KSM. Einiges weniger, nämlich 27.3% bevorzugten dabei das KSM.

5.7 Subjektive Usability von Augmented Chemistry

Die Boxplots in Abbildung 5.28 bilden die Verteilungen der Skalen des SUMI ab.

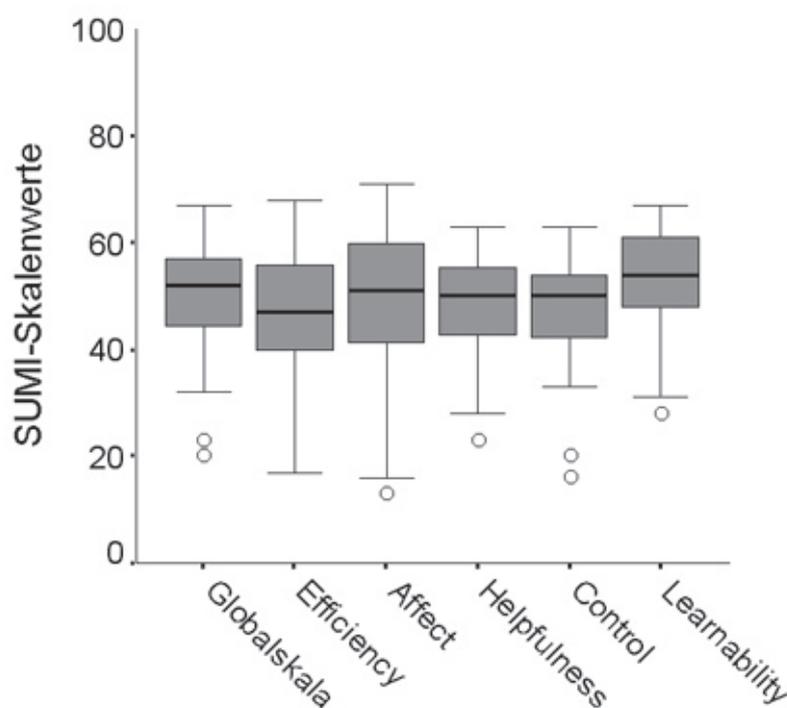


Abb. 5.28: Verteilungen der SUMI-Skalenwerte (N = 23).

Alle Mediane bewegen sich in einem sehr ähnlichen Bereich zwischen 47 und 54 Punkten. Die Werte streuen sehr stark nach unten und Werte über 60 werden kaum erreicht. Die Mediane befinden sich alle nahe dem Mittelwert der Skala, was bedeutet, dass AC im Schnitt in den einzelnen Usability-Kriterien mittelmässig beurteilt wurde. Bei den Dimensionen *affect* und *efficiency* wurde AC vereinzelt auch sehr schlecht bewertet. Es gab also durchaus auch Personen, welche emotional sehr negativ auf AC reagierten und das Ausmass an Unterstützung durch das System bei der Aufgabenbearbeitung als sehr gering einschätzten. Eine geringe Effizienz-Bewertung kann auf eine inkonsistente und befremdliche Funktionsweise der Software zurückgeführt werden.

Die Resultate aus dem SUMI zeigen, dass die Usability von AC durch die Benutzer nicht zufriedenstellend beurteilt wurde. Die Werte bewegen sich mehrheitlich um 50 Punkte und deuten somit auf eine eher geringe Usability von AC hin. In den einzelnen Dimensionen erhielt AC zwar zumeist eine Bewertung oberhalb von 40 Punkten (Werte kleiner als 40 deuten auf einen Verbesserungsbedarf hin), dennoch finden sich durchwegs noch knapp 25% aller Werte unterhalb dieser Grenze. Da sich bei den einzelnen Skalen ein sehr ähnliches Bild zeigt, wird AC in allen erhobenen Dimensionen ein Verbesserungspotenzial zugesprochen. Von einer "guten" Software kann man erst ab Werten über 60 sprechen, was bei AC nur in wenigen Fällen erreicht wurde.

Nun könnten die einzelnen Items von SUMI noch genauer analysiert werden, indem die jeweiligen Antwortprofile mit den aus der Normstichprobe erwarteten Profilen verglichen würden. Die Übereinstimmung zwischen den beobachteten und erwarteten Werten würde mehr Aufschluss darüber geben, worin Probleme oder positive Seiten bei AC liegen. Da die einzelnen Items aber sehr oft mit 'unentschieden' beantwortet wurden, ergibt sich aus ihnen kein klares Bild für eine Interpretation. Folglich wird hier auf diese Analyse verzichtet.

5.8 Zusammenhänge zwischen Lernerfolg und subjektiven Variablen

Im Folgenden wird untersucht, ob und wie die Variablen zur Lernwirksamkeit (praktische Problemlösefähigkeit und Lernerfolg) mit den subjektiven Variablen emotionale Stimmung (Aktiviertheit, Valenz, Dominanz) und Beanspruchung zusammenhängen. Die Korrelationen werden dabei getrennt für die Lernsituation mit AC und derjenigen mit dem KSM analysiert, um herauszufinden, ob und wie sich die Zusammenhänge zwischen den beiden Lernsituationen unterscheiden. Für die allgemeinen Zusammenhänge sei auf die Korrelationsmatrix in Tabelle 5.1 in Kapitel 5.1 verwiesen.

Aus Tabelle 5.12 und 5.13 ist ersichtlich, dass die meisten Korrelationen in der AC-Lernsituation stärker ausfallen als in der KSM-Lernsituation. Bei den Korrelationen der Problemlösefähigkeit mit den drei Dimensionen zur emotionalen Stimmung finden sich bei AC die stärksten Zusammenhänge bei der Valenz ($r = .46$, $p < .05$) und der Dominanz ($r = .50$, $p < .05$). Diese Korrelationen fallen in der KSM-Lernsituation deutlich geringer aus ($r = .23$ bzw. $r = .07$, $p > .05$). Die positiven Korrelationen bedeuten, dass je besser die Aufgaben mit dem Lernsystem gelöst werden konnten, desto eher wurde bei den Upn ein positives Gefühl ausgelöst, und desto eher wurde die Situation danach unter Kontrolle empfunden. Die Tatsache, dass die Korrelationen bei der AC-Lernsituation stärker ausgeprägt sind als in der KSM-Lernsituation, könnte damit zusammenhängen, dass die Qualität des Arbeitens mit dem Lernsystem bei AC

stärker gewichtet wird. So ruft eine gute praktische Problemlösefähigkeit beim Lernen mit AC eher eine angenehme Stimmung hervor als beim Lernen mit dem KSM.

Tab. 5.12: Lernsituation AC: Korrelationen der objektiven Variablen zur Lernwirksamkeit mit den subjektiven Variablen emotionale Stimmung und Beanspruchung.

	I	II	III	IV	V	VI
I Problemlösefähigkeit	-					
II Lernerfolg	.14	-				
III Aktiviertheit (SAM)	-.21	-.29	-			
IV Valenz (SAM)	.46*	.24	-.31	-		
V Dominanz (SAM)	.50*	.10	-.10	.28	-	
VI Beanspruchung (NASA)	-.69**	-.32	.34	-.59**	-.53**	-

(Korrelationen nach Spearman)

Anmerkungen: n = 23. *, **: p<.05, p<.01

Tab. 5.13: Lernsituation KSM: Korrelationen der objektiven Variablen zur Lernwirksamkeit mit den subjektiven Variablen emotionale Stimmung und Beanspruchung.

	I	II	III	IV	V	VI
I Problemlösefähigkeit	-					
II Lernerfolg	-.53**	-				
III Aktiviertheit (SAM)	-.26	.12	-			
IV Valenz (SAM)	.23	-.24	.08	-		
V Dominanz (SAM)	.07	-.15	-.23	.15	-	
VI Beanspruchung (NASA)	-.02	-.22	.15	-.52*	-.01	-

(Korrelationen nach Spearman)

Anmerkungen: n = 23. *, **: p<.05, p<.01

Bei der Aktiviertheit zeigen sich in beiden Lernsituationen vergleichbare Zusammenhänge ($r = -.21$ bei AC; $r = -.26$ beim KSM, $p > .05$). Die negativen Korrelationen bedeuten: Je besser die Lernaufgaben gelöst werden konnten, desto eher führte dies zu einer niedrigen Aktiviertheit. Was die subjektive Beanspruchung betrifft, so steht diese beim Lernen mit AC in deutlich stärkstem Zusammenhang mit der praktischen Problemlösefähigkeit ($r = -.69$, $p < .01$). In diesem Zusammenhang zeigt sich auch der grösste Unterschied zwischen den beiden Lernsituationen. Je geringer die Beanspruchung durch das Lernen mit AC empfunden wurde, desto besser wurden die Lernaufgaben mit AC gelöst. Die entsprechend geringe Korrelation in der KSM-Lerneinheit weist darauf hin, dass die Beanspruchung beim Lernen mit dem KSM keine so grosse Rolle gespielt hat für die Qualität der Aufgabenbearbeitung ($r = -.02$, $p > .05$).

Beim Lernerfolg liegen die Korrelationen mit den subjektiven Variablen in beiden Lernsituation in ähnlicher Grössenordnung vor. Auch mit dem Lernerfolg steht die Beanspruchung in der AC-Lernsituation in stärkstem Zusammenhang ($r = -.32, p > .05$). Beim KSM hängt die Beanspruchung schwach bis mittelstark mit dem Lernerfolg zusammen ($r = -.22, p > .05$). Auffallend sind die beiden entgegengesetzten Korrelationen der Valenz mit dem Lernerfolg in beiden Lernsituationen. Die mittelstarken Zusammenhänge sind bei AC positiv ($r = .24, p > .05$), beim KSM hingegen negativ ausgeprägt ($r = -.24, p > .05$). Beim Lernen mit AC bedeutet dies: Je positiver das durch das Lernen verursachte Gefühl war, desto eher wurden im Kurztest die besseren Resultate erzielt. Beim Lernen mit dem KSM verhält es sich genau umgekehrt.

Des Weiteren steht die Beanspruchung beim Lernen mit AC in einem starken Zusammenhang ($r = -.53, p < .01$) mit dem Dominanz erleben, während sich beim KSM entsprechend ein sehr schwacher Zusammenhang zeigt ($r = -.01, p > .05$). Je geringer folglich die Beanspruchung durch das Lernen mit AC empfunden wurde, desto eher berichteten die Lernenden anschliessend, die Situation unter Kontrolle zu haben. Beim Lernen mit dem KSM hat die Beanspruchung durch das Lernsystem aber keinen Einfluss auf das Dominanzempfinden.

Bei beiden Lernsituationen steht die Valenz in einem stark negativen Zusammenhang mit der Beanspruchung (AC: $r = -.59, p < .01$; KSM: $r = -.52, p < .05$). Dies bedeutet, dass je geringer die empfundene Beanspruchung durch die Lernsituation war, desto positivere Gefühle wurden bei den Upn ausgelöst.

Insgesamt liegen in der AC-Lernsituation also mehrheitlich stärkere Zusammenhänge vor als in der KSM-Lernsituation. Bei der praktischen Problemlösefähigkeit kann man folglich darauf schliessen, dass die Qualität des Arbeitens mit dem Lernsystem bei AC stärker gewichtet wird und sie somit das emotionale Empfinden stärker beeinflusst als dies beim KSM der Fall ist. Eine zunehmende Beanspruchung wirkt sich zudem nur negativ auf das Lernen mit AC nicht aber auf das Lernen mit dem KSM aus.

5.9 Prädiktive Bedeutung von Lernmethode, Periode, emotionaler Stimmung und Beanspruchung für den Lernerfolg

In einem letzten Schritt wird noch die Bedeutung der Lernmethode, der Periode und der subjektiven Variablen emotionale Stimmung und Beanspruchung als Determinanten des Lernerfolgs mittels einer multiplen Regression untersucht (Tabelle 5.14).

Tab. 5.14: Ergebnisse der multiplen Regression des Lernerfolgs auf die Variablen Lernmethode, Periode, emotionale Stimmung und Beanspruchung.

Variablen	B	Beta	r	sr ¹	p
Lernmethode	5.04	.10	.11	.18	= .25
Periode	-40.23	-.77	-.81	-.83	< .01
Aktiviertheit	-.90	-.07	-.14	-.12	= .44
Valenz	-2.63	-.17	.08	-.26	= .11
Dominanz	-.37	-.03	.08	-.05	= .75
Beanspruchung	-.53	-.34	-.36	-.44	< .01

Anmerkungen: n = 48, R = .86, R² = .75, F = 19.09, df = 6/39, p < .01

¹ Semipartialkorrelationen: widerspiegeln Grösse der spezifischen Beiträge der Prädiktoren

Lernmethode, Periode, emotionale Stimmung und Beanspruchung erklären zusammen 75% der Varianz des Lernerfolgs. Dieser Prognosewert ist sehr hoch, und es zeigt sich, dass der Erklärungsbedarf hauptsächlich durch den Prädiktor Periode gedeckt wird. 69% der Varianz des Lernerfolgs wird allein durch die Periode erklärt, gefolgt von 20% erklärter Varianz durch die Beanspruchung. Der interessanteste Prädiktor, die Lernmethode, vermag allein nur 3% der Varianz des Lernerfolgs zu erklären. Der Vergleich der standardisierten Regressionskoeffizienten (Beta-Gewichte) zeigt, dass die Periode mit dem grössten Gewicht in die Vorhersage des Lernerfolgs einfließt (Beta = -.77). Die prädiktive Bedeutung der Beanspruchung ist etwas kleiner (Beta = -.34). Beide Koeffizienten sind signifikant (p < .01). Eine noch kleinere Rolle spielt die Lernmethode (Beta = .10). Nur die Periode und die Beanspruchung leisten aber einen signifikanten (p < .01) Beitrag für die Vorhersage des Lernerfolgs. Die Lernmethode sollte folglich kein zu grosses Gewicht in der Vorhersage des Lernerfolgs erhalten, wenn man die erheblich grössere Bedeutung der Periode und der Beanspruchung in Betracht zieht.

Inhaltlich gesehen erstaunt die dominierende Rolle der Periode. Verschiedene Lerninhalte können den Lernerfolg demnach sehr stark beeinflussen. Mit welchem Lernsystem dieser Lerninhalt dabei erarbeitet wird, zeigt sich in vorliegender Studie für den Lernerfolg als praktisch irrelevant.

6. Diskussion

In der vorliegenden Arbeit ging es um eine Evaluation von Augmented Chemistry, einem neuartigen computergestützten Lernsystem zum Lernen von Molekülen. Dabei interessierte, ob Augmented Chemistry (AC) hinsichtlich der Lernwirksamkeit und Benutzerakzeptanz vergleichbar oder sogar besser ist als das Kugel-Stab-Modell (KSM), ein herkömmlich verwendetes Lernmittel. Von neuen, insbesondere computergestützten, Lernsystemen wird erwartet, dass die Schüler besser lernen und motivierter sind. Ob dies durch den Einsatz von AC erreicht werden kann, wurde versucht mit vorliegender Arbeit zu beantworten. Um die Lernwirksamkeit und die Akzeptanz der beiden Lernsysteme vergleichend zu testen, wurde ein AB/BA-Design verwendet. Im Folgenden wird zuerst eine Bewertung und Interpretation der Ergebnisse vorgenommen. Dem schliesst sich eine genauere Analyse des Potenzials von AC als erfolgreiche Lernumgebung an. Nach einer kritischen Reflektion eigener Untersuchungsansätze werden zum Schluss praktische Konsequenzen für AC diskutiert und Vorschläge für weiterführende Studien entwickelt.

6.1 Interpretation der Ergebnisse

In den folgenden Kapiteln werden die einzelnen Ergebnisse zu den untersuchten Variablen im Hinblick auf die Fragestellungen gedeutet und versucht miteinander in Beziehung zu setzen.

6.1.1 Lernwirksamkeit

Während mit der praktischen Problemlösefähigkeit das Arbeiten mit AC und dem KSM an sich bewertet wurde, so wurde mit dem Lernerfolg das Lernergebnis, sprich die Effektivität der Lernsysteme, erhoben. Das KSM dient in dieser Studie als Referenz. Allfällige Unterschiede zwischen den Lernmethoden sollen andeuten, welche Lernmethode besser ist; keine Unterschiede weisen auf Gleichwertigkeit der beiden Lernmethoden hin, was somit positiv bewertet wird.

Die Ergebnisse der Untersuchung haben gezeigt, dass Personen, die mit dem KSM gelernt haben, die bessere praktische Problemlösefähigkeit aufweisen als Personen, die mit AC gelernt haben. Die Aufgaben wurden folglich mit dem KSM signifikant besser gelöst als mit AC. Im Lernerfolg zeigte sich hingegen kein bedeutender Unterschied zwischen den beiden Lernmethoden.

Die Ergebnisse zur praktischen Problemlösefähigkeit lassen darauf schliessen, dass das KSM die bessere Unterstützung beim Erarbeiten von Lerninhalten bietet als AC. Diesen Ergebnissen

zufolge wäre für diese Funktion das herkömmliche Lernsystem KSM gegenüber dem neuen AC vorzuziehen. Es werden im Folgenden Erklärungen für das bessere Abschneiden des KSM bei den Lernaufgaben gesucht. Es lag die nicht optimale Tatsache vor, dass die Versuchsbedingungen für die beiden Lernmethoden nicht vollständig identisch waren. Bei der Aufgabenbearbeitung mit dem KSM gab es einige Vorteile gegenüber AC, was sich wiederum auf das positive Resultat beim KSM ausgewirkt haben mag. Während bei AC die Lernenden die Aufgaben völlig selbständig, ohne fremde Hilfe lösen mussten, so war eine KSM-Lerneinheit von der Interaktion mit einer Lehrperson abhängig. Beim KSM wurden die Lerntexte erst ausgehändigt, nachdem die Lehrperson das gebaute Molekül für richtig beurteilt hatte. Dadurch war es bis zu diesem Punkt mit dem KSM also praktisch unmöglich, Fehler zu machen. Bei AC hingegen wurden die Lerntexte automatisch eingespielt, nachdem die Moleküle fertig konstruiert waren. Ob das gebaute Molekül dabei richtig oder falsch war, mussten die Lernenden selbst beurteilen. Die beim KSM gewährleistete Kontrollmöglichkeit über die Richtigkeit des gebauten Moleküls fehlte somit bei AC. Mit AC mussten die Lernenden folglich aufmerksamer und selbständiger arbeiten als mit dem KSM. Dieses völlig selbständige Aneignen von Wissen bei AC war für die Lernenden aber teilweise ungewohnt und befremdend. Hier wäre zu klären, ob diese Form der selbständigen Wissensaneignung mit einer Ergänzung des Unterrichts durch einen Medieneinsatz von dem betreffenden Lehrinstitut gewünscht wird. Eine veränderte Einstellung zum Lernen wäre dann auf Seite der Lernenden notwendig, und ihre Selbständigkeit und Eigenverantwortung müsste unterstützt werden (Strittmatter & Mauel, 1995). Dabei sollte der Lehrer eine unterstützende und nicht mehr nur vermittelnde Funktion wahrnehmen (Bauer, 1995).

Das Ausmass der praktischen Problemlösefähigkeit sollte auch aussagen, wie gut der Umgang mit dem jeweiligen Lernsystem war. Ob nun die schlechtere Problemlösefähigkeit beim Lernen mit AC auf geringere Kompetenzen im Umgang mit dem System schliessen lässt, soll im Weiteren beantwortet werden. Aus Beobachtungen der AC-Lerneinheiten geht hervor, dass bei der Konstruktion von Molekülen mit AC gelegentlich die Bindungsart zu wenig beachtet wurde. So passierte es in wenigen Fällen, dass aus Versehen Doppelbindungen anstatt Einfachbindungen gesetzt wurden. Dieser Fehler konnte beim KSM nicht passieren, da Doppelbindungen nicht durch die Länge der Stäbchen symbolisiert sind, sondern tatsächlich durch zwei Stäbchen dargestellt werden. Die genauen Mechanismen bei AC zur Auswahl der Bindungsart über das Tetraedermodell (siehe Kapitel 2.5.3, Abbildung 2.10) wurden in der Einführung in die Handhabung von AC zwar erläutert, aber wahrscheinlich nicht nachhaltig. Dass die Einführung zum Untersuchungszeitpunkt für die Lernenden bereits zwei bis drei Wochen zurücklag, war auch nicht förderlich für das Arbeiten mit AC. Im Umgang mit den Bindungsarten gab es also offensichtlich noch Probleme. Das Vorliegen von Defiziten in der Bedienung und Handhabung von AC wurde auch durch die Präferenzbefragung bestätigt. Hier zeigt sich, wie wichtig eine genaue Kenntnis im Umgang und auch eine Vertrautheit mit dem System ist, damit es seine Wir-

kung entfalten kann. Man muss aber nicht nur auf der Lernerseite ansetzen und die Fähigkeit der Benutzer im Umgang mit AC verbessern, auch auf Seite von AC sind Verbesserungen in Bezug auf eine Vereinfachung der Bedienung notwendig. Zudem müssen auch die technischen Mängel bei AC behoben werden, welche in Kapitel 5.5.2 unter Punkt b) beschrieben wurden.

Im Gegensatz zu den Lernaufgaben zeigten sich beim Lernerfolg keine bedeutenden Unterschiede zwischen den Lernmethoden. Es ist erstaunlich, dass sich die Resultate in den Lernaufgaben nicht im Lernerfolg niederschlagen. Ein besseres Resultat in den Lernaufgaben sollte nämlich auch zu einem besseren Resultat im Kurztest führen. Die widersprüchlichen Resultate könnten damit zusammenhängen, dass das angeeignete Wissen in den Lernaufgaben zu wenig verfestigt werden konnte, da der Kurztest unmittelbar an eine Lerneinheit folgte. Folglich kann die Qualität der bearbeiteten Aufgaben hier nicht als ein Mass dafür herangezogen werden, wie gut der Lerninhalt gelernt wurde. Anhand der Problemlösefähigkeit können aber Rückschlüsse getroffen werden, wie gut der Umgang mit dem jeweiligen Lernsystem beurteilt werden kann.

Betrachtet man die beiden Untersuchungszeitpunkte, so fällt beim Lernerfolg der starke Periodeneffekt auf. Die Tatsache, dass zum zweiten Untersuchungszeitpunkt deutlich schlechtere Resultate im Kurztest erzielt wurden, könnte auf die unterschiedlichen Lerninhalte und die etwas komplexere Version des Kurztests zum zweiten Untersuchungszeitpunkt zurückgeführt werden. Im Kurztest beim zweiten Versuch mussten im Vergleich zum ersten Kurztest komplexere Moleküle und Molekülnamen reproduziert werden, sowie Inhalte aus den Lerntexten erinnert werden, was im ersten Versuch nicht verlangt wurde. Abgesehen davon kann aber konstatiert werden, dass die neue Lernumgebung AC bezüglich des Lernerfolgs zwar nicht wie erhofft besser dafür aber genauso förderlich ist wie die herkömmliche Lernumgebung KSM. Dieses Resultat bestätigt die vielen Befunde aus evaluativen Einzelstudien, dass computergestützte Lernmedien gegenüber konventionellen Medien nicht effektiver sind (vgl. Hasebrook, 1995; Kerres, 2001; Schulmeister, 1997). Bisherige Untersuchungen haben gezeigt, dass computergestützte Lernmedien zwar über Potenziale zur Verbesserung der Lernleistung verfügen, dennoch würden sich nur wenige oder gar keine positiven Auswirkungen auf den Lernerfolg zeigen (Hasebrook, 1995). Das Ziel liegt nun darin, herauszuarbeiten wie diese Potenziale von AC das Lernen positiv beeinflussen könnten.

6.1.2 Beanspruchung

Weder beim gewichteten Gesamtwert noch bei den einzelnen Subskalen des Nasa Task Load Index liessen sich signifikante Unterschiede zwischen den beiden Lernmethoden aufzeigen. Eine Ausnahme bildet die Subskala geistige Beanspruchung, bei welcher die Lernenden mit AC signifikant höhere Werte erreichten als mit dem KSM. Dies bedeutet, dass sich die Lernenden beim Arbeiten mit AC geistig beanspruchter fühlten als beim Arbeiten mit dem KSM. AC erforderte demnach mehr geistige Anstrengung bei der Informationsaufnahme und Informationsverarbeitung (z.B. Denken, Entscheiden, Erinnern, Hinsehen, Suchen), sowie eine höhere Genauigkeit. In dieser Hinsicht wurde das Arbeiten mit AC als anspruchsvoller und komplexer empfunden als das Arbeiten mit dem KSM. Diese höhere Beanspruchung durch die AC-Lernsituation könnte eine Erklärung für das schlechtere Abschneiden in den Aufgaben mit AC bieten. Ein Grund für die höhere geistige Beanspruchung beim Arbeiten mit AC wird darin gesehen, dass das indirekte Zusammenbauen von Molekülen mit Buch, Greifer und Würfel komplexer ist und somit mehr an Denkarbeit erfordert als das direkte Hantieren mit Kugeln und Stäbchen beim KSM. Diese Erklärung ist wiederum ein Rückschluss auf die Bedienung und heisst soviel wie, dass die Lernenden bei AC durch die Bedienung mehr in Anspruch genommen werden als beim KSM.

6.1.3 Emotionale Reaktion

Die emotionale Reaktion auf die beiden Lernsysteme wurde über die drei Dimensionen Aktiviertheit, Valenz und Dominanz ermittelt. Am schwierigsten sind die Resultate zur Aktiviertheit zu interpretieren. Grund ist der Interaktionseffekt zwischen der Lernmethode und der Periode. Eine Interpretation des Behandlungseffekts aus dem AB/BA-Design ist in diesem Fall zwar nicht möglich, es wird aber versucht die beobachtete wachsende Tendenz in der Aktiviertheit bei AC und die sinkende Tendenz beim KSM vom ersten zum zweiten Untersuchungszeitpunkt zu interpretieren. Im Sinne eines Carry-Over-Effekts könnte es durchaus sein, dass die durch das Lernsystem ausgelöste Aktiviertheit im zweiten Versuch von der emotionalen Reaktion auf das im ersten Versuch getestete Lernsystem abhängig war. Die Beurteilung im zweiten Versuch wäre dann nicht unabhängig vom ersten Versuch gewesen. Beim signifikanten Interaktionseffekt zwischen Lernmethode und Periode könnte es auch sein, dass der Haupteffekt der Gruppe dieses signifikante Resultat bewirkt hat (siehe dazu Kapitel 4.6.1, Tabelle 4.4). Zwar sollten durch die Parallelisierung der Gruppen allfällige Gruppeneffekte möglichst ausgeschlossen werden können, da es aber eine Vielzahl an Lernervariablen gibt, welche niemals alle berücksichtigt und kontrolliert werden können, sind solche Gruppeneffekte aber nicht unwahrscheinlich. Betrachtet man die Resultate zur Aktiviertheit für beide Gruppen, so fällt auf, dass bei beiden Gruppen die Aktiviertheit nach dem Lernen mit AC höher war als nach dem Lernen

mit dem KSM. Die zweite Gruppe reagierte durchgehend mit erregungsbezogeneren Gefühlen als die erste Gruppe. Ob das signifikante Resultat als Interaktionseffekt oder Gruppeneffekt gedeutet wird, ändert nichts an der Tatsache, dass das Gesamtergebnis, in Bezug auf einen Behandlungseffekt, nicht eindeutig interpretiert werden kann, da hierfür sowohl Carry-Over- als auch Gruppeneffekte ausgeschlossen werden müssten.

Die Valenzdimension entspricht einem Mass für Wohlbefinden (vgl. Schallberger, 1999). Mit der Valenz wird das Ausmass der Freude gemessen, wie angenehm oder unangenehm ein Gefühl gerade ist. Bei der Freude handelt es sich um ein Merkmal intrinsischer Motivation. Die Resultate zur Valenz fielen in vorliegender Studie gut aus. Das Arbeiten mit AC löste vorwiegend positive Gefühle aus, was die Akzeptanz dieses Systems wiederum positiv beeinflusst. Im Vergleich zum KSM schnitt AC zwar tendenziell schlechter ab, signifikante Unterschiede konnten aber nicht gefunden werden. Hinsichtlich der Akzeptanz liegt das neue Lernsystem AC also etwa gleichauf mit dem altbewährten KSM. Nun ist die positive Wirkung der wahrgenommenen Freude auf die Systemakzeptanz und die entscheidende Bedeutung dieser Empfindung für die Entstehung von intrinsischer Motivation unbestritten. Die Freude ist aber nur ein Teil des gesamten emotionalen Erlebens eines Benutzers. Welche Emotionen und dazugehörigen Systemeigenschaften für ein positives Systemerleben und schliesslich für die Systemakzeptanz wirklich von Bedeutung sind, bleibt unklar (Trautmann, 2004). Dies müsste in weiteren empirischen Untersuchungen genauer überprüft werden.

Die Valenz tritt immer in Kombination mit einem bestimmten Ausmass an Aktiviertheit auf. Demnach ist eine Differenzierung eines angenehmen oder unangenehmen Gefühls nach dem Grad der Aktiviertheit möglich. Eine solche Differenzierung der Erfassung des Befindens wurde in vorliegender Arbeit nicht vorgenommen, sollte es in zukünftigen Untersuchungen mit den beiden emotionalen Dimensionen Valenz und Aktiviertheit aber gemacht werden. Beschränkt man sich auf die Extremausprägungen der Aktiviertheits- und Valenzdimension, so lassen sich vier Zustände als Kombination der beiden Dimensionen definieren (vgl. Schallberger, 1999, S. 18): Eine hohe Valenz, also ein positives Gefühl bei hoher Aktiviertheit, entspricht etwa einem Zustand, der durch Begeisterung und überschäumende Freude gekennzeichnet ist. Ein positives Gefühl bei niedriger Aktiviertheit wäre ein Zustand völliger Entspannung, verbunden mit innerer Ruhe und Ausgeglichenheit. Ein negatives Gefühl (niedrige Valenz) verbunden mit hoher Aktiviertheit ergeben Ärger oder Angst, verbunden mit niedriger Aktiviertheit Langeweile (Abbildung 6.1). Analysen von Emotions- und Stimmungsbegriffen zeigen, dass sich das gesamte Spektrum an möglichen Gefühlen in der in Abbildung 6.1 dargestellten Ebene kreisförmig anordnen lässt, in einem sogenannten Circumplex-Modell (vgl. Schallberger, 1999). Die Circumplex-Struktur von Befindlichkeitsbegriffen wurde von Russell (1980; zit. nach Schallberger, 1999, S. 19) untersucht. Eine gleichzeitige Berücksichtigung der beiden Dimensionen

Valenz und Aktiviertheit würde zu einer deutlichen Differenzierung in der Erfassung des Befindens führen. Genauere Erkenntnisse über die Auswirkungen von AC auf die Emotionen bei den Benutzern wären die Folge.

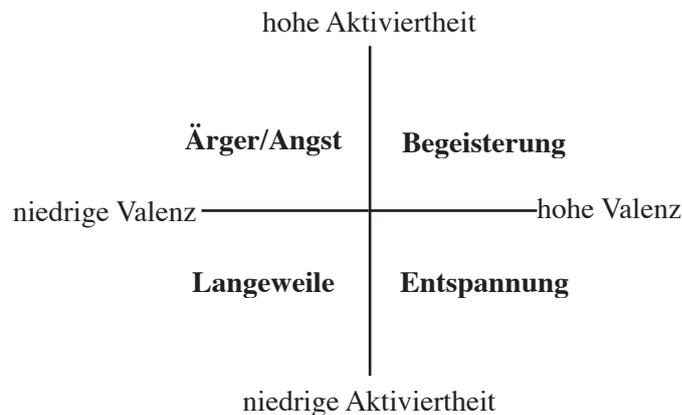


Abb. 6.1: Befindlichkeiten aus der Kombination der Extremausprägungen in der Valenz- und Aktiviertheitsdimension (abgeleitet aus der Circumplex-Struktur der Emotionen nach Russell (1980)).

6.1.4 Präferenzbeurteilung

Die Resultate aus der Präferenzbefragung sind sehr aufschlussreich. Die Ergebnisse zeigen, dass bei vier Aspekten AC und bei weiteren vier Aspekten das KSM bevorzugt wurde. Bei den deutlichen AC-Präferenzen handelt es sich eher um emotionale Aspekte (Spassfaktor, Absicht einer zukünftigen Anwendung) und um für ein solches Lernsystem zur Zweckerfüllung massgebende Kriterien (bessere Veranschaulichung, besserer Zugang zum Lerninhalt). AC hat den Lernenden folglich mehr Spass gemacht und ihnen einen besseren Zugang zum Lerninhalt ermöglicht, vielleicht gerade durch eine bessere Veranschaulichung. Somit wurde ein wichtiges Ziel von AC erreicht. Wie aus den Antworten auf die offenen Fragen hervorgeht, beurteilten viele Probanden am Lernen mit AC das Arbeiten am Computer als sehr positiv. Auch die Technologie bei diesem System faszinierte einige. Dabei spielt es anfänglich wahrscheinlich gar keine so grosse Rolle, welche Inhalte das computergestützte Lernprogramm bietet. Der Computer an sich wird als das attraktivere Lernmedium betrachtet. Das Lernen am Computer kann also durchaus eine motivationsfördernde Wirkung haben. Für das Lernen langfristig förderlich ist das aber nicht. Es besteht die Gefahr eines Neuigkeitseffekts, wonach eine motivationssteigernde Wirkung eines Lernsystems mit der Zeit wieder verschwindet (Kerres, 2001, Schulmeister, 1997).

Bei den KSM-Präferenzen standen Aspekte der Bedienung und Handhabung des Lernsystems im Vordergrund. Dies verdeutlicht, dass AC prioritär im Bereich Bedienung an Verbesserung

bedarf. Wie bereits die Ergebnisse zur Problemlösefähigkeit und zur geistigen Beanspruchung gezeigt haben, gab es noch Probleme im Umgang mit AC, was das Lernen wiederum negativ beeinflusst. Worin genau die Probleme im Umgang mit AC lagen, lässt sich teilweise aus den Antworten auf die offen formulierten Fragen lesen. So wurde von den Probanden insbesondere der Greifer als störend empfunden, das System reagierte nicht immer richtig und auch die Wahl der Bindungsart bereitete vielen Schwierigkeiten.

6.1.5 Zufriedenheit

Die Zufriedenheit der Probanden wurde durch einen eigens konstruierten Fragebogen erfasst. Über die Bewertung verschiedener Item-Aussagen, Antworten aus offenen Fragen und einer Benotung der beiden Lernsysteme sollten Aussagen in Bezug zur Zufriedenheit getroffen werden. Ein Teil der Antworten auf die offenen Fragen wurde bereits zur Interpretation der Ergebnisse aus der Präferenzbefragung in Kapitel 6.1.4 herangezogen. Die Analyse der Item-Antworten hat gezeigt, dass sich die Lernenden mit AC und jene mit dem KSM hinsichtlich ihrer Zufriedenheit nur sehr gering voneinander unterscheiden. Mit beiden Lernmethoden waren die Lernenden zufrieden bis sehr zufrieden. Die Items wurden aber beim KSM durchwegs leicht besser bewertet als bei AC, was bedeutet, dass das Lernen mit dem KSM bei den Probanden mehr Zufriedenheit ausgelöst hat als das Lernen mit AC. Am deutlichsten mehr Zustimmung gaben Personen, die mit dem KSM gelernt hatten, bei der Bewertung der Unterstützungsfunktion und der Hilfe des Lernsystems zu einem besseren Verständnis der Lerninhalte. Diese Befunde können durch die Präferenzbeurteilung bestätigt werden, bei welcher dem KSM die bessere Unterstützung beim Lernen zugesprochen wurde. Die Tatsache, dass die Aufgaben mit dem KSM bedeutend besser gelöst wurden, erklärt das Zustandekommen dieses Beurteilungsmusters. Eine lernfördernde Wirkung der Lernsysteme wurde bei beiden Lernmethoden gleichermassen eher negativ bewertet. Deutlich bessere Bewertungen erhielten hingegen emotionale oder motivationale Faktoren wie der Spass am Lernen oder ob eine zukünftige Anwendung der Lernmethode wünschenswert sei. Dieser Umstand, dass trotz einer relativ gering eingeschätzten Lernwirksamkeit der Lernmethoden die emotionale Bewertung eher positiv ausfiel, könnte auf einen Neuigkeitseffekt deuten. Dieser Neuigkeitseffekt kann man neben AC auch dem KSM zuschreiben, denn auch das KSM wurde zum Untersuchungszeitpunkt erstmals im Unterricht eingesetzt.

Die Bewertung der beiden Lernmethoden auf einer Notenskala sollte schliesslich ein allgemeineres Mass für die Zufriedenheit der Probanden mit den Lernmethoden liefern. Auch hier fielen die Unterschiede nicht ins Gewicht, wobei erneut das KSM die leicht bessere Bewertung erhielt, ein Muster, das sich bereits durch die gesamten Ergebnisse der Studie gezeigt hat.

6.1.6 Akzeptanz: Eine Integration aller untersuchten Aspekte

Damit AC sowohl von Lernenden als auch von Lehrenden als Lernsystem akzeptiert werden kann, spielen alle in dieser Studie untersuchten Aspekte eine Rolle. AC kann nur Akzeptanz finden, wenn es lernwirksam ist, wenn die Benutzer emotional positiv auf das System reagieren, wenn sie in Bezug auf ihre Beanspruchung nicht über- oder unterfordert und wenn sie allgemein zufrieden mit der Lernsituation sind. Ob AC diese Kriterien erfüllt, kann nur in Bezug auf die Referenzmethode KSM beurteilt werden. Es wird davon ausgegangen, dass das altbewährte KSM sowohl bei den Lehrkräften als auch bei den Lernenden gut akzeptiert ist. AC müsste in den untersuchten Aspekten also zumindest vergleichbare Resultate wie das KSM liefern, um ebenfalls Akzeptanz zu finden. Ein zukünftiger Einsatz wäre aber nur durch ein besseres Abschneiden von AC gerechtfertigt. Es hat sich nun gezeigt, dass AC hinsichtlich der für die Akzeptanz relevanten Variablen nicht besser war als das KSM. Tendenziell schnitt das KSM sogar leicht besser ab als AC.

Freude als emotionale Reaktion auf ein Programm trägt deshalb zur Akzeptanz bei, weil sie als intrinsischer Motivator gilt (Davies et al., 1992; zit. nach Trautmann, 2004, S. 23). Auch Spass am Lernen ist ein intrinsischer Motivator und sollte die Aneignung von Lerninhalten erleichtern. Draper (1999; zit. nach Trautmann, 2004, S. 23) betont die Wichtigkeit von Spass und Spiel, welche nicht nur auf Computerspiele beschränkt werden sollen. Spass am Lernen wurde in dieser Arbeit mehrmals erhoben, einerseits im Fragebogen zur Zufriedenheit (Items 1 und 6: “Das Arbeiten mit AC/KSM hat Spass gemacht” und “Mit AC/KSM macht das Lernen mehr Spass”), andererseits anhand der Präferenzbeurteilung. Da der Spassfaktor mehrfach erhoben wurde, sollen die verschiedenen Ergebnisse miteinander verglichen werden. Erstaunlicherweise zeigen sich widersprüchliche Befunde: Bei der Präferenzbeurteilung gibt es ein eindeutiges Resultat zu Gunsten von AC, im Fragebogen zur Zufriedenheit aber, wo die Beurteilungen zum Spassfaktor ohne Referenzbezug vorgenommen wurden, zeigt sich eine Tendenz in Richtung KSM. Die Unterschiede sind im letzteren Fall aber gering und der Zufall kann aufgrund der nicht signifikanten Ergebnisse nicht ausgeschlossen werden. Mit beiden Lernsystemen machte das Lernen grundsätzlich Spass. Geht es aber darum, sich in einem direkten Vergleich für ein Lernsystem zu entscheiden, so kam heraus, dass das Lernen mit AC deutlich mehr Spass gemacht hatte als das Lernen mit dem KSM.

Allgemein kann festgehalten werden, dass bei den inferenzstatistischen Verfahren meistens nur sehr kleine und kaum signifikante Unterschiede zwischen den beiden Lernmethoden gefunden werden konnten. Dieses wenig differenzierte Bild macht es schwierig die genauen Vorteile von AC gegenüber dem KSM zum Vorschein zu bringen. Die geringen Unterschiede lassen zumindest auf eine Gleichwertigkeit der beiden Lernmethoden im Lernerfolg und der Akzeptanz schliessen, was erfreulich ist. Die Ergebnisse zeigen aber, dass für einen erfolgreichen Einsatz

von AC ein erheblicher Mehrwert gegenüber dem KSM erreicht werden muss. Worin dieser Mehrwert liegt, wird im Kapitel 6.4 genauer analysiert.

6.2 Zusammenhänge

Betrachtet man die Zusammenhänge zwischen den subjektiven Variablen und der Lernwirksamkeit getrennt nach den beiden Lernmethoden, so zeigen sich teilweise deutlich stärkere Zusammenhänge in der AC-Lernsituation. Am aufschlussreichsten sind die Zusammenhänge zwischen der Problemlösefähigkeit und der Valenz, der Dominanz und der Beanspruchung. Bei den emotionalen Dimensionen bedeuten die Resultate, dass je besser die Aufgaben gelöst werden konnten desto angenehmere Gefühle wurden bei den Lernenden ausgelöst und desto eher hatten sie das Gefühl, die Situation unter Kontrolle zu haben. Dass diese Zusammenhänge bei AC stärker ausfallen als beim KSM mag bedeuten, dass die Qualität des Arbeitens mit dem Lernsystem bei AC stärker gewichtet wird als beim KSM. Gerade um angenehme Gefühle auszulösen, spielt es bei AC demnach die grössere Rolle als beim KSM, wie gut die Aufgabebearbeitung funktioniert. Da die Valenz wiederum für die Systemakzeptanz bedeutend ist, kann man schliessen, dass die Akzeptanz von AC stark davon abhängt, wie gut die Aufgaben mit dem System gelöst werden können. Da der Zusammenhang beim KSM in deutlich geringerem Ausmass besteht, ist die Akzeptanz vom KSM folglich weniger aufgabengebunden. Eine positive Erfahrung bei der Auseinandersetzung mit dem Lernsystem ist bei AC demzufolge viel wichtiger für die Systemakzeptanz als beim KSM. Diese stärkere Gewichtung bei AC könnte damit zusammenhängen, dass AC im Vergleich zum KSM etwas Neues und noch Unbekanntes ist, und negative Erfahrungen im Umgang mit Unbekanntem wahrscheinlich eher zu negativen Gefühlen führen als bei Bekanntem wie dem KSM. Bei der Beanspruchung wirkt sich eine zunehmende Beanspruchung nur negativ auf die Leistung in Lernaufgaben bei AC nicht aber beim KSM aus. Aus diesen Resultaten kann man schliessen, dass Lernende vor dem Einsatz von AC ausreichend vertraut mit dem System gemacht werden müssen. Erst wenn Unsicherheiten bei den Benutzern beseitigt sind, kann AC seine Wirkung als Lernhilfe entfalten.

Beim Lernerfolg sind die (schwachen) Zusammenhänge mit den subjektiven Variablen bei beiden Lernmethoden sehr ähnlich. Aus der Regressionsanalyse geht auch hervor, dass die Lernmethode ein eher unbedeutender Prädiktor für die Vorhersage des Lernerfolgs ist. Einen weit- aus grösseren Erklärungswert weist die Periode auf, aber auch der Beanspruchung kommt eine bedeutende Rolle für den Lernerfolg zu. Dies zeigt, dass der Einfluss der Lernmethode nicht überbewertet werden darf; unterschiedliche Lerninhalte oder die empfundene Beanspruchung kommen stärker zum Tragen. Lernerfolg darf man folglich nicht wie so oft an der alleinigen Gestaltung von Lernumgebungen festmachen. Lernfördernde Potenziale von Medien entfalten

sich nämlich nicht von selbst. Lernende müssen diese vielmehr selber erschliessen unter bestimmten kognitiven, motivationalen und situativen Bedingungen Tergan (1997).

6.3 Subjektive Usability von Augmented Chemistry

Die Usability von AC wurde von den Lernenden im Schnitt mittelmässig eingeschätzt. Daraus geht hervor, dass die Lernenden AC zwar mehrheitlich als brauchbar betrachteten, von einer ausserordentlich guten Software kann man aufgrund der Daten aber nicht ausgehen. Was die einzelnen Usability-Kriterien anbelangt, so zeigen sich kaum Unterschiede in der Beurteilung der einzelnen Dimensionen. Die einzelnen Dimensionen werden im Folgenden interpretiert.

Effizienz (Efficiency): Das eher gering empfundene Ausmass an Unterstützung, das die Benutzer bei ihrer Arbeit durch AC erfahren haben, konnte auch durch die Präferenzbefragung bestätigt werden. Mehr Probanden empfanden das KSM als die bessere Unterstützung zum Lernen als AC. Dem Usability-Resultat zufolge ist eine effektivere, schnellere und ökonomischere Aufgabendurchführung mit AC anzustreben.

Emotionale Bewertung (Affect): Dieser emotionale Aspekt stellt das Nutzererleben in den Vordergrund. Die Daten weisen ebenfalls auf einen Verbesserungsbedarf von AC bezüglich dieser Dimension hin. Für eine Systemoptimierung müsste aber zuerst eine einheitliche Vorstellung darüber vorliegen, welche Konzepte massgeblich für ein positives Systemerleben sind.

Selbstbeschreibungsfähigkeit (Helpfulness): Auch die Forderung nach einer Verständlichkeit des Produkts aus sich selbst heraus und nach dem Angebot situationsunabhängiger Hilfen wurde nicht zufriedenstellend erfüllt. Dies wäre aber eine Voraussetzung für intuitives Lernen. Bei AC ist keine Hilfefunktion im System integriert, was bei einer Systemoptimierung zur Verbesserung der Benutzerfreundlichkeit mitberücksichtigt werden könnte.

Erlernbarkeit (Learnability): Der benötigte Aufwand, den Umgang mit AC zu erlernen, wurde am besten von allen Dimensionen bewertet. Die Probanden hatten das Gefühl mit relativ wenig Aufwand, AC zu beherrschen. Die Tatsache, dass bei den Versuchen aber doch einige Probleme im Umgang mit AC vorlagen, könnte damit zusammenhängen, dass zwischen dem Erlernen der Benutzung von AC und dem Untersuchungszeitpunkt zwei bis drei Wochen lagen. Probleme beim Arbeiten mit AC könnten somit durch die relativ lange Nichtbenutzung des Systems aufgetreten sein. Benutzer sollten aber auch nach langer Nichtbenutzung von Programmen schnell wieder mit dem System arbeiten können. Durch eine Vereinfachung der Benutzerschnittstelle von AC könnte diese Form der Benutzerfreundlichkeit besser erreicht werden.

Steuerbarkeit (*Control*): Diese Dimension wurde im Vergleich zu den übrigen relativ gut beurteilt. AC reagierte gemäss den Lernenden relativ konsistent auf ihre Eingaben. Die Bedingung, dass der Benutzer die Software und nicht die Software ihn unter Kontrolle haben soll, wird für AC als gegeben betrachtet. Der Aspekt der Kontrolle über ein System wurde auch bei der emotionalen Stimmung über die Dominanzdimension erfasst. Die Resultate dazu zeigen ebenfalls, dass die Probanden die Situation nach dem Lernen mit AC gut unter Kontrolle hatten.

Die einzelnen Items aus dem SUMI waren teilweise schwer auf AC anzuwenden, was die Erfassung der Usability von AC durch diese Methode in Frage stellt. Viele Items wurden bevorzugt mit 'unentschieden' beantwortet, was - im Falle der Absicht ein milderes Urteil abzugeben - die Resultate negativer ausfallen lässt. Aussagen über die Usability von AC sollten folglich nicht nur anhand der verwendeten Methode gemacht werden. Vielmehr sollten in zukünftigen Studien weitere Methoden zur Usability-Messung verwendet werden, beispielsweise die Methode des Lauten Denkens (*Thinking Aloud*) (vgl. Hegner, 2003). Hierbei müsste der Benutzer während der Interaktion mit AC all seine Gedanken, Gefühle und Meinungen äussern. Diese Technik kann auch ohne konkrete Aufgabenstellung angewandt werden. Mit dieser Messmethode können schnell Probleme mit dem System aufgedeckt werden, und eine Klärung, warum allfällige Fehler aufgetreten sind, ist möglich. Neben der Technik des Lauten Denkens sollten für die Bewertung der Usability von AC aber auch die entstandenen log files ausgewertet oder Interviews durchgeführt werden, um ein differenzierteres Bild über die Usability von AC zu erhalten.

6.4 Potenzial von Augmented Chemistry als erfolgreiche Lernumgebung

Aufgrund der Resultate dieser Studie kann festgehalten werden, dass AC sowohl in der Lernwirksamkeit als auch in der Benutzerakzeptanz vergleichbar mit dem KSM ist. Zwar schnitt AC bei der praktischen Problemlösefähigkeit signifikant schlechter ab als das KSM, was sich aber durch die vorhandenen Probleme bei der Bedienung, die mangelhafte Benutzerfreundlichkeit und technische Mängel (siehe Kapitel 5.5.2) von AC erklären lässt.

Diese Studie hat gezeigt, dass eine innovative Lernform - wie es AC ist - ein bereits bekanntes Lernmittel wie das KSM nicht so einfach ersetzen kann. Das KSM erfüllt seinen Zweck, es ist einfach und sowohl den Lehrpersonen als auch den Lernenden gut vertraut. Geht es rein um das Konstruieren von Molekülmodellen ist der Vorteil vom KSM unbestritten. Das Problem bei AC könnte deshalb bei der Benutzerschnittstelle liegen. Die Eingabewerkzeuge sind komplex und für die Konstruktion von Molekülen zu wenig intuitiv. Im Prinzip wäre das KSM die perfekte Schnittstelle. Eine Konstruktion der Moleküle bei AC nach dem Vorbild vom KSM könnte Vor-

teile bringen. Über die Molekülkonstruktion hinaus beginnt klar der Mehrwert von AC, welcher durch die Interaktivität und Multimedialität geboten wird. So kann man bei AC verschiedene Darstellungsformen der Moleküle wählen, die Molekülnamen werden automatisch eingeblendet, Dipolmomente können angezeigt werden, etc. In diesen zusätzlichen Funktionen, welche aus AC mehr als ein reines Molekülbauprogramm machen, liegt die Chance für AC, sich als besseres Lernsystem im Unterricht zu etablieren. Zuerst muss aber die Hürde zu einer einfacheren Schnittstelle überwunden werden, um dann den bestehenden Mehrwert von AC ausnutzen zu können oder noch zusätzlichen zu schaffen. Dieser Mehrwert muss aber von den zukünftigen Benutzern des Systems, oder von solchen, die es einsetzen möchten, auch gewünscht werden. Nur so kann eine Marktlücke geschlossen werden und AC zum Erfolg verhelfen. Mit AC wurde zwar ein gutes Produkt entwickelt, in Zukunft müsste man sich aber bereits bei der Entwicklung von solchen Systemen intensiv auch mit didaktischen Fragen und möglichen Einsatzszenarien auseinandersetzen. Dies wurde bei AC zu wenig berücksichtigt.

Für einen erfolgreichen Einsatz von AC im Chemieunterricht spielen neben der in vorliegender Studie untersuchten Lernwirksamkeit, Akzeptanz und Brauchbarkeit auch Faktoren wie die curriculare Einbindung (vgl. Glowalla und Häfele, 1995), didaktische Aspekte und der Interaktivitätsgrad eine wichtige Rolle. Im Folgenden wird kurz auf die drei letztgenannten Faktoren eingegangen.

6.4.1 Curriculare Einbindung von Augmented Chemistry

AC kann sowohl von den Lernenden als auch von den Lehrenden eingesetzt werden. Bei der Bereitstellung für die Lehrenden kann AC primär zur Unterstützung der Instruktion eingesetzt werden, beispielsweise als Präsentationshilfe oder Anschauungsmaterial im Rahmen von Präsenzveranstaltungen. Dazu müsste das computergenerierte Bild über einen Beamer projiziert werden. Auf Seite der Lernenden können mit AC Lernmaterialien oder Werkzeuge zum selbstgesteuerten Lernen zur Verfügung gestellt werden. AC könnte in einzelne Veranstaltungsstunden eingebunden werden und zwar als Mittel zur Veranschaulichung und Verfestigung von Lerninhalten, welche im traditionellen Unterricht vermittelt wurden. Sinnvoll, aber aus Kostengründen nicht einfach zu realisieren, wäre die Ausrüstung der Schulen mit ausreichend vielen und leistungsfähigen Notebooks, auf welchen AC lauffähig ist. Kopfhörer für die auditiv präsentierten Lerntexte sind unumgänglich. Um vernünftig mit Schülern an AC arbeiten zu können, ist eine relativ lange Einarbeitungszeit der Schüler und auch der Lehrer in das System notwendig, was aber Zeit kostet. Auf jeden Fall müssen für die curriculare Einbindung jeweils die Besonderheiten der Lehr- und Lernsituation, des organisatorischen Umfelds, der Zielgruppe, der technischen Infrastruktur sowie der softwaretechnischen Lösung berücksichtigt werden (Blumstengel, 1998).

6.4.2 Didaktische Aspekte von Augmented Chemistry

Die Konzeption von AC erlaubt - je nach Anwendung - verschiedene Lernmodelle einzubeziehen. Grundsätzlich basiert AC auf einer konstruktivistischen Lernauffassung, bei der die Selbstkonstruktion von Wissen durch aktives Ausprobieren bedeutend ist (Schulmeister et al., 2003). Für einen zielgerichteten und effizienten Lernprozess ist auch die Instruktion wichtig. Diese instruktionale Komponente wird aber nicht automatisch durch AC gewährleistet, dazu müssen zuerst konkrete Aufgaben durch die Lehrperson konstruiert werden. Mit diesen Aufgaben können die Lernenden dann mit Hilfe von AC ihr Wissen aneignen. Diese Notwendigkeit der Entwicklung von Lernaufgaben als instruktionale Komponente bedeutet für die Lehrperson zwar einiges an Aufwand, macht den Einsatz von AC aber flexibler. AC lässt sich nämlich so durch die didaktische Intention der Lehrperson ergänzen und dem Vorwissen der Lernenden anpassen.

Neben der Funktion eines Lernsystems, mit dem die Lernenden sich selbständig Wissen über den Aufbau und die Struktur von Molekülen aneignen können, kann AC auch als reines Visualisierungsprogramm verwendet werden.

6.4.3 Interaktivitätsgrad von Augmented Chemistry

Bei AC stellt sich nicht die Frage, ob es interaktiv ist oder nicht, es geht vielmehr darum zu fragen *wie* interaktiv AC ist.

Bei AC spielt der Interaktionsstil der direkten Manipulation eine Rolle. Die Merkmale von direkter Manipulation liegen in (vgl. Shneiderman, 1997):

- der ständigen Darstellung der relevanten Objekte und Aktionen auf dem Bildschirm.
- der Möglichkeit der direkten physischen Manipulation der Objekte.
- der Ausführung inkrementeller und reversibler Operationen, deren Ergebnis in Echtzeit sichtbar wird.

Für AC werden diese drei Kriterien als erfüllt betrachtet. Die Möglichkeit, dass die Benutzer durch direkte Manipulation gestaltend eingreifen können, erhöht den Interaktivitätsgrad von AC. Die Benutzer können durch eigenes Handeln in den Lernprozess eingreifen, indem sie unter anderem nach Belieben den Darstellungsmodus der Moleküle verändern oder ein gewünschtes Molekül in der Bibliothek suchen können. So sind sie in der Lage, ihren Lernweg selbst zu bestimmen und auf diesem Weg ihr Wissen zu vertiefen. Der Benutzer ist nicht mehr nur Rezipient, sondern ist in den Lernprozess gestaltend einbezogen. Ein aktiver und planvoller Umgang mit dem System, welches Werkzeugcharakter hat, steht im Vordergrund (vgl. Baum-

gartner & Payr, 1994). Dieses Mass an Eingriffs- und Steuerungsmöglichkeiten, welches dem Benutzer gewährt wird, ist ein wichtiges Merkmal von Interaktivität. Die genannten Eigenschaften weisen auf einen hohen Interaktivitätsgrad von AC hin. Trotzdem könnte das System durch Weiterentwicklungen noch interaktiver gemacht werden, um somit weiteren Mehrwert zu schaffen. Je stärker sich Lernende nämlich in eine Lernanwendung einbringen können, desto attraktiver wird die Anwendung und motiviert auch zum Lernen (vgl. Aufenanger, 1999). Die Interaktivität könnte beispielsweise dadurch gesteigert werden, dass von AC selbst Aufgaben generiert würden, dass Moleküle miteinander in Reaktion gebracht oder funktionelle Gruppen in Molekülen hervorgehoben werden könnten.

6.5 Probleme und kritische Betrachtungen

Die Studie wurde im natürlichen Umfeld einer Berufsschule während der regulären Unterrichtszeit der Upn durchgeführt. Dies ist einerseits ein grosser Vorteil, da natürliche Bedingungen eine Generalisierung der Ergebnisse auf ähnliche Situationen zulässt. Die externe Validität ist also gegeben. Andererseits ist eine natürliche Versuchsumgebung auch mit Nachteilen verbunden. So musste die Versuchsplanung den Gegebenheiten von *aprentas* angepasst werden. Ausfälle des Schulbetriebs und infolgedessen Unterbrüche im Versuchsablauf mussten in Kauf genommen werden. Die Versuche mussten während der regulären Unterrichtszeit der Probanden durchgeführt werden, was eine äusserst knapp bemessene Zeit für die Überprüfung des Lernerfolgs zur Folge hatte.

Eine bedeutende Einschränkung für die Aussagekraft der Studie liegt in der geringen Nutzungsdauer der Lernsysteme von lediglich 45 Minuten (vgl. Aufenanger, 1999). Noch problematischer ist die äusserst knapp bemessene Zeit von 10 Minuten für den Kurztest. Es können so nur beschränkt Aussagen über Lern- und Nutzerverhalten, geschweige denn über den Lernerfolg gemacht werden. Zudem stellt die Kontrolle des Lernerfolgs lediglich durch ein Testen des deklarativen und prozeduralen Faktenwissens in einem zunehmend konstruktivistisch geprägten Lernverständnis ein weiteres Problem dar (Kerres, 2001). Nach Baumgartner (1999) geht es bei interaktiven Lernprogrammen darum, dass Schüler den Lerngegenstand erforschen und sich in konstruktiver Weise Wissen aneignen können. Anstatt Wissen durch Reproduktionsleistung zu testen, sollten vermehrt *Lernprozesse* beobachtet und bewertet werden (vgl. Kozma, 1994; zit. nach Freudenreich & Schulte, 2002, S. 2). Eine Analyse des durch log files aufgezeichneten Benutzerverhaltens ist in zukünftigen Studien erforderlich. Lernerfolgskontrollen unter dem Paradigma des konstruktivistischen Lernens erweisen sich aber als problematisch, da hier der Erwerb komplexer, individuell unterschiedlicher Fertigkeiten geprüft werden soll (Schaumburg & Rittmann, 2000).

Ein weiterer kritischer Aspekt liegt bei der Überprüfung des Mehrwerts von AC. In dieser Studie wurden im Prinzip nur Eigenschaften von AC getestet, welche auch für das KSM in irgendeiner Weise überprüft werden konnten. Dies war notwendig, damit die Vergleichbarkeit der beiden Lernmethoden gewährleistet war. Die Frage ist nun, wie man einen Mehrwert von AC empirisch überprüfen kann, damit die Vergleichbarkeit gewährleistet bleibt. Die Konzentration hätte sich mehr auf die Vorteile von AC gegenüber dem KSM richten sollen (z.B. auf die verschiedenen Darstellungsmöglichkeiten der Moleküle), und ob diese Funktionen, welche das KSM nicht bieten kann, Lerneffekte zu steigern vermögen. Das KSM wurde für diese Studie - mit nicht unbedeutendem Aufwand - sehr an AC angepasst, so dass sich die Besonderheiten von AC nicht mehr hervorheben konnten. Dies könnte für die geringen und statistisch kaum signifikanten Unterschiede zwischen den beiden Lernmethoden verantwortlich sein.

Des Weiteren wurden in dieser Arbeit allein Lernsituationen miteinander verglichen und dadurch zu wenig spezifische Lernervoraussetzungen und Einstellungen berücksichtigt. Dies führt nach Carey (2001) zwangsläufig zu nicht-signifikanten Ergebnissen. Es gibt eine Vielzahl an Lernervariablen, welche alle zu kontrollieren ein schier unmögliches Unterfangen darstellt. Der Einfluss einiger weniger Lernervariablen wurde durch parallelisierte Verteilung der Lernenden auf die beiden Lernsituationen versucht konstant zu halten, was aber wiederum zu einer "Nivellierung" der Unterschiede in der Wirkung der unabhängigen Variablen führen kann (Schulmeister et al., 2003, S. 5). Dies hat sich in vorliegender Arbeit bestätigt, indem kaum signifikante Resultate gefunden werden konnten. Pädagogisch interessante Ergebnisse hätten durch eine Abgrenzung der beiden Lernsituation hinsichtlich didaktischer Aspekte (z.B. Ausmass der Einwirkungsmöglichkeiten durch den Benutzer oder Ausmass der Selbststeuerungsmöglichkeit des Lernens) und durch eine Gruppierung der diesen Lernsituationen ausgesetzten Personen in Bezug auf spezifische Lernervariablen erlangt werden können. So hätten die Probanden beispielsweise nach ihrer Motivation (extrinsisch versus intrinsisch), nach ihrem Vorwissen (viel versus wenig) oder nach verschiedenen Lernstrategien in Gruppen eingeteilt werden können. Differenziertere Aussagen über die Wirkung von AC wären die Folge.

6.6 Ausblick

Eine Behebung technischer Mängel und eine Vereinfachung der Bedienung wären die ersten Schritte, AC zu einem effektiveren und brauchbareren Lernsystem für den Chemieunterricht zu machen. Hierbei sind die speziellen Eingabewerkzeuge - vor allem der Würfel und der Greifer - zu überdenken, um sie durch intuitivere Objekte zu ersetzen. Dazu müsste aber die gesamte Schnittstelle neu überarbeitet und auch andere Steuerungsmöglichkeiten überlegt werden, also wie man beispielsweise einfacher die Bindungsart wählen oder Atome auswählen kann.

Um einen bedeutenden Mehrwert gegenüber dem KSM zu schaffen, müsste eine Reihe zusätzlicher Funktionen bereitgestellt werden, wie die Möglichkeit, zwei Moleküle gleichzeitig darzustellen und diese in Reaktion miteinander zu bringen, oder aber das Hervorheben funktioneller Gruppen in einem gebauten Molekül. So wie AC zur Zeit besteht, sind die Chancen gering, dass es an Schulen eingesetzt wird. Für die Weiterentwicklung von AC ist ein Austausch mit Lehrinstituten, welche AC einsetzen könnten, notwendig. Die genauen Bedürfnisse, sowohl von Lehrkräften als auch von Lernenden, müssen genauer diskutiert werden. Diese Interaktion zwischen Entwicklern und Praktikern könnte iterativ zu einem besseren AC und Einsatzszenarium führen. Eine Umstellung des erst auf Linux laufenden Systems auf Mac- und Windows-Lauffähigkeit ist in Aussicht. Weitere Systemoptimierungen hinsichtlich der Bedienung sind ebenfalls geplant.

Gerade die Akzeptanz eines Lernsystems wird auch stark von technischen Elementen beeinflusst. So kann ein Lernprogramm aufgrund eines Mangels als praktisch unbrauchbar eingestuft werden, obwohl dieses Problem vermutlich innerhalb relativ kurzer Zeit behoben werden könnte. Zudem können innovative Lösungen zunächst zu schlechteren Lernergebnissen führen, weil sie anfangs ungewohnt sind. Dieser Effekt kann sich aber bereits nach kurzer Zeit umkehren (vgl. Blumstengel, 1998). Dies soll verdeutlichen, dass eine Lernumgebung wie AC nach den vorliegenden ernüchternden Ergebnissen nicht voreilig als unbrauchbar abgetan werden sollte. Eine sorgfältige Abwägung des Aufwands und der Kosten, die bestehenden Mängel zu beheben, sollte stattdessen vorgenommen werden.

Im Folgenden werden Vorschläge für weiterführende Studien entwickelt.

- Zwei verschiedene Medientypen - einen virtuellen und einen realen Typ - miteinander zu vergleichen, könnte problematisch sein. Es wäre sinnvoll, weitere Studien innerhalb desselben Medientyps durchzuführen, also AC mit einem anderen interaktiven, computergestützten Lernmedium zu vergleichen.
- Im verwendeten Versuchsmodell konnten nur kurzzeitige Lerneffekte erfasst werden. In weiteren Studien sollte man möglichst auch Langzeitlerneffekte untersuchen, um aussagen zu können, wie dauerhaft das Lernen ist.
- In den Versuchen wurden auch Daten erhoben, welche im Rahmen dieser Arbeit nicht ausgewertet wurden. So wurde bei AC das Benutzerverhalten mit log files aufgezeichnet und entsprechend beim KSM Videoaufzeichnungen der Lerneinheiten gemacht. Um dem Anspruch gerecht zu werden, bei interaktiven Medien Lernerfolg auch durch die Beobachtung und Bewertung von Lernprozessen zu beurteilen (Kozma, 1994; zit. nach Freudenreich & Schulte, 2002, S. 2), wäre in zukünftigen Untersuchungen der Einbezug solcher log files sinnvoll.

- Bei AC müsste in zukünftigen Untersuchungen auch die multimediale Komponente genauer untersucht werden, insbesondere die Kombination verschiedener Modalitäten (visuell/akustisch) und Codierungen (verbal/nonverbal). AC bietet nämlich nicht nur visuell nonverbale (bildliche), sondern auch akustisch verbale und nonverbale Informationen in Form von gesprochenen Lerntexten oder Geräuschen. Die Relevanz dieser Kombinationen sollte noch genauer untersucht werden, auch ob und wie eine multimediale Lernumgebung die Aneignung von Wissen fördern kann. Es wäre sinnvoll zu klären, über welche Modalitäten und Codierungen die Inhalte auf optimale Weise transportiert werden können. Die Wirkung kombinierter Darbietung von visuellen und auditiven Informationen wurde bislang noch nicht ausreichend untersucht. Es herrscht also noch Forschungsbedarf.

Bei der Weiterentwicklung von AC müssen neben softwaretechnischen und gestalterischen Parametern in Zukunft vermehrt auch didaktische Parameter berücksichtigt werden. Eine Ergänzung der technischen Perspektive mit einer pädagogisch-psychologischen Forschungsperspektive würde zu neuen Erkenntnissen über die wechselseitige Abhängigkeit von Wissenserwerb und Gestaltung von computergestützten Lernumgebungen führen (vgl. Schwan & Buder, 2002, S. 128). Neben der bis anhin einseitigen Konzentration auf die Entwicklung von AC als optimales Lernprogramm sollte auch vermehrt die pädagogische Einbettung mitberücksichtigt werden (Aufenanger, 1999). Erst wenn man weiss, wie und welchen Lerninhalt man mit AC am besten lernt und unter welchen Bedingungen es an Schulen eingesetzt werden muss, kann das Lernen mit AC zum Erfolg führen. Ausserdem braucht es Lehrpersonen, die von AC voll und ganz überzeugt sind, es ausreichend beherrschen, um es dann auch sinnvoll im Unterricht einsetzen zu können.

7. Zusammenfassung

In der vorliegenden Arbeit wurde ein innovatives computergestütztes Lernsystem für den Chemieunterricht, mit dem Namen Augmented Chemistry, hinsichtlich der Lernwirksamkeit und Benutzerakzeptanz evaluiert. Die Evaluation fand im Umfeld einer Berufsschule statt und wurde in Form einer Vergleichsstudie durchgeführt. Dabei wurde die neue Lernmethode Augmented Chemistry (AC) mit einem herkömmlich verwendeten Lernmittel, dem Kugel-Stab-Modell (KSM) verglichen. Die Stichprobe umfasste 26 Personen zwischen 15 und 22 Jahren. Es waren Biologielaboranten im ersten Ausbildungsjahr an der Berufsschule *aprentas* in Muttenz. *aprentas* ist ein Ausbildungszentrum für technische und naturwissenschaftliche Berufe, insbesondere in der chemischen und pharmazeutischen Industrie. AC wurde also durch potenzielle Endbenutzer getestet.

Die Versuche waren nach einem AB/BA-Design aufgebaut. Danach wurden die beiden Lernmethoden zu zwei Zeitpunkten getestet und zwar anhand von zwei Lerngruppen, welche in unterschiedlicher Reihenfolge mit den Lernsystemen arbeiteten. Die Probanden mussten Aufgaben mit den Systemen lösen und ihr so erworbenes Wissen anschliessend in einem kurzen Test unter Beweis stellen. Subjektive Masse wie die empfundene Beanspruchung beim Lernen, die emotionale Reaktion auf die Lernsysteme, die Zufriedenheit mit der Lernsituation, die subjektive Usability von AC und die Präferenz für eine Lernmethode wurden mit Fragebogen erhoben.

Es folgen nun die wichtigsten Resultate zu den fünf Fragen der vorliegenden Lizentiatsarbeit.

In einer ersten Frage wurde überprüft, ob sich die beiden Lernmethoden, AC und das KSM, hinsichtlich ihrer Lernwirksamkeit unterscheiden. Hierbei wurde die praktische Problemlösefähigkeit und der Lernerfolg untersucht, welche über die Qualität der gelösten Aufgaben mit den Systemen und über einen Kurztest operationalisiert wurden. Die Ergebnisse zeigten, dass sich Personen, welche mit AC, und solche, die mit dem KSM gearbeitet hatten, signifikant in ihrer Problemlösefähigkeit unterschieden: Die Aufgaben wurden mit dem KSM bedeutend besser gelöst als mit AC. Im Lernerfolg konnten dagegen keine signifikanten Unterschiede zwischen den beiden Lernmethoden gefunden werden.

In einer zweiten Frage ging es um die Unterschiede zwischen AC und dem KSM hinsichtlich subjektiver Masse bei den Benutzern, anhand derer Aussagen über die Benutzerakzeptanz gemacht werden sollen. Weder bei der Gesamtbeanspruchung, noch bei der emotionalen Stimmung, noch bei der Zufriedenheit konnte bei den Probanden ein signifikanter und somit bedeutender Unterschied gefunden werden. Tendenziell war die Beanspruchung aber beim Arbeiten mit dem KSM höher als bei AC, und das Arbeiten mit dem KSM löste etwas angenehmere Gefühle aus als das Arbeiten mit AC. Einzig bei einer Subdimension der Beanspruchung - bei der geistigen Beanspruchung - konnte ein signifikanter Unterschied gefunden werden. Beim

Arbeiten mit AC fühlten sich die Probanden geistig bedeutend beanspruchter als beim Arbeiten mit dem KSM. Bei der Präferenzbeurteilung zeigte sich, dass in emotionaler und inhaltlicher Hinsicht AC bevorzugt wurde, während bei Aspekten der Bedienung und Handhabung der Lernsysteme die Präferenz beim KSM lag. So hatte das Arbeiten mit AC den Lernenden zwar mehr Spass gemacht und konnte ihnen den Lerninhalt besser veranschaulichen, das KSM fanden die meisten jedoch einfacher zu bedienen und angenehmer im Umgang.

In einer dritten Frage wurde der Zusammenhang zwischen der Lernwirksamkeit und den subjektiven Variablen Beanspruchung und emotionale Stimmung untersucht. Die Zusammenhänge wurden dabei getrennt nach den beiden Lernsituationen analysiert. Die meisten Zusammenhänge fielen in der AC-Lernsituation stärker aus als in der KSM-Lerneinheit. Bei AC konnten mittlere und signifikante Zusammenhänge zwischen der praktischen Problemlösefähigkeit und den emotionalen Dimensionen Valenz und Dominanz gefunden werden: Je besser die Aufgaben gelöst werden konnten, desto angenehmere Gefühle wurden bei den Lernenden ausgelöst, und desto eher hatten sie das Gefühl, die Situation unter Kontrolle zu haben. Diese Zusammenhänge fielen beim KSM schwach bis sehr schwach aus, was bedeutet, dass die durch das KSM ausgelöste emotionale Stimmung nur wenig von der Qualität der Aufgabenbearbeitung abhängt. Bei AC konnte auch ein mittlerer bis starker und signifikanter Zusammenhang zwischen der praktischen Problemlösefähigkeit und der Beanspruchung gefunden werden: Je höher die Beanspruchung beim Arbeiten mit AC empfunden wurde, desto schlechter wurden auch die Aufgaben gelöst. Dieser beim KSM sehr schwach ausfallende Zusammenhang deutet darauf hin, dass sich eine zunehmende Beanspruchung beim KSM im Gegensatz zu AC nicht negativ auf die Aufgabenbearbeitung auswirkt.

Die vierte Frage setzte sich mit der subjektiven Usability von AC auseinander. Die Usability wurde von den Lernenden nur mittelmässig beurteilt. Alle fünf Usability-Kriterien (*affect, control, efficiency, helpfulness, learnability*) erhielten eine ähnlich Beurteilung. Aufgrund der Resultate müssten alle fünf Kriterien für AC nochmals überdacht und Verbesserungsmöglichkeiten diesbezüglich ausgearbeitet werden.

In einem letzten Schritt wurde versucht, anhand der angefallenen Ergebnisse zu beantworten, worin das Potenzial von AC als erfolgreiche und wirksame Lernumgebung liegt. Die Resultate haben gezeigt, dass AC weder in der Lernwirksamkeit noch in der Benutzerakzeptanz besser ist als das herkömmliche KSM. Bedeutende Unterschiede zwischen den beiden Lernmethoden wurden kaum gefunden. Dies bestätigt die zahlreichen Einzelbefunde aus Evaluationsstudien, welche belegen konnten, dass computergestützte Lernmedien gegenüber traditionellen Medien nicht effektiver sind (Hasebrook, 1995; Kerres, 2001; Schulmeister, 1997). Es zeigte sich, dass die Aufgaben mit dem alten Lernmedium KSM sogar signifikant besser gelöst werden konnten als mit dem neuen AC. Das KSM war dadurch das bessere Lernmittel, um die gegebenen Lern-

inhalte zu erarbeiten. Das Problem bei AC wurde im Speziellen bei der Benutzerschnittstelle gesehen. Die verschiedenen Eingabewerkzeuge sind sehr komplex und zu wenig intuitiv, um Molekülmodelle problemlos zu konstruieren, was sich auf die Aufgabenbearbeitung auswirkt. Als reines Molekülbauprogramm bietet AC keinen Vorteil gegenüber dem KSM. Das Konstruieren von Molekülmodellen müsste bei AC stärker nach dem Vorbild vom KSM erfolgen. Das Potenzial von AC findet sich in einer Reihe von zusätzlichen Funktionen, über welche das KSM nicht verfügt. Dieser Mehrwert, welcher über die Möglichkeit der Interaktivität und Multimedialität von AC gewährleistet wird, sollte in Zukunft noch ausgebaut werden. Auch müsste das vorhandene Potenzial von AC in Zukunft genauer untersucht werden, inwiefern dadurch das Lernen wirklich positiv beeinflusst werden könnte.

Eine Optimierung von AC in verschiedenen Punkten ist Voraussetzung für einen Erfolg dieses Lernsystem. Durch die Beseitigung technischer Mängel, die Vereinfachung der Bedienung und der Bereitstellung einer Reihe zusätzlicher Funktionen könnte ein deutlicherer Mehrwert gegenüber dem KSM geschaffen werden. Im Weiteren müssen neben dieser Konzentration auf Entwicklungsmöglichkeit auch vermehrt didaktische Fragen und die pädagogische Einbettung von AC berücksichtigt werden (vgl. Aufenanger, 1999). Dazu ist aber eine Zusammenarbeit zwischen Programmentwicklern und Pädagogen wünschenswert, um letztendlich den Bedürfnissen der Endbenutzer von AC vollständig gerecht zu werden.

8. Literaturverzeichnis

- Alsdorf C. & Bannwart E. (1995). Virtuelle Realität: Erfahrbare Informationen im Cyberspace. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 437-450). Weinheim: Psychologie-Verl.-Union.
- Aufenanger, S. (1999). Lernen mit neuen Medien - Was bringt es wirklich? Forschungsergebnisse und Lernphilosophien. *Zeitschrift für Medienpädagogik*, 23 (Heft 4), 4-8.
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. (2001). Recent Advances in Augmented Reality. *IEEE Computer Graphics & Applications*, 34-47.
- Bauer, W. (1995). Multimedia in der Schule? In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 377-399). Weinheim: Psychologie-Verl.-Union.
- Baumgartner, P. & Payr, S. (1994). *Lernen mit Software*. Innsbruck: Österreichischer Studienverlag.
- Baumgartner, P. (1999). Evaluation mediengestützten Lernens: Theorie - Logik - Modelle. In M. Kindt (Hrsg.), *Projektelevaluation in der Lehre: Multimedia an Hochschulen zeigt Profil(e)* (S. 63-99). Münster: Waxmann.
- Bente, G., Krämer, N. C. & Petersen, A. (2002). Virtuelle Realität als Gegenstand und Methode der Psychologie. In G. Bente, N. C. Krämer & A. Petersen (Hrsg.), *Virtuelle Realitäten* (S. 1-31). Göttingen: Hogrefe.
- Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal*, 4, 115-150. [On-line]. Available: <http://www.usability.serco.com/papers/qusab95.pdf>
- Bevan, N. (1999). Quality in Use: Meeting User Needs for Quality. *Journal of System and Software*. [On-line]. Available: <http://www.usability.serco.com/papers/qiuse.pdf>
- Biocca, F. & Delaney, B. (1995). The vision of virtual reality. In F. Biocca & M. Levy (Eds.), *Communication in the age of virtual reality* (pp. 57-126). Hillsdale: Erlbaum.
- Blumstengel, A. (1998). *Entwicklung Hypermedialer Lernsysteme*. Dissertation. [On-line]. Available: <http://dsor.upb.de/de/forschung/publikationen/blumstengel-diss/>
- Boling, E. & Sousa, G. (1993). Interface design issues in the future of business training. *Business Horizons*, 36, 54.
- Bonner, J. (1988). Implications of cognitive theory for instructional design. *Educational Communication and Technology Journal*, 36, 3-14.
- Bowman, D., Wineman, J. & Hodges, L. (1999). The educational value of an information-rich virtual environment. *Presence*, 8, 317-331.

-
- Bruner, J. (1961). The act of discovery. *Harvard Educational Review*, 31, 21-32.
- Büchner, R. (1995). Computer ist mehr. In: Computer ist mehr - Multimedia und Schule - Ein Fortbildungsmodell der Akademie für Lehrerfortbildung Dillingen. München: Manz.
- Carey, J. M. (2001). *Effective Student Outcomes: A Comparison of Online and Face-to-Face Delivery Modes*. [On-line]. Available: http://www.ed.psu.edu/acsde/deos/deosnews/deosnews11_9.asp
- Chang, C. Y. (2000). Enhancing tenth graders' earth science learning through computer-assisted instruction. *Journal of Geoscience Education*, 48, 636-640.
- Chang, C. Y. (2001). A problem-solving based computer-assisted tutorial for the earth sciences. *Journal of Computer Assisted Learning*, 17, 263-274.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition). Hillsdale: Erlbaum.
- Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, 22, 1111-1132.
- Dede, C., Salzman, M. & Loftin, R. (1996). *Science Space: Virtual realities for learning complex and abstract scientific concepts*. Proceedings of IEEE Virtual Reality Annual International Symposium (pp. 246-253). Hillsdale: Erlbaum.
- Degenhardt (2001). Möglichkeiten empirischer Erfassung der Computernutzung von Schüler/innen im Unterricht. *Online-Zeitschrift Medienpädagogik*. [On-line]. Available: <http://www.medienpaed.com/00-2/degenhardt1.pdf>
- Dillon, A. (2001). Beyond usability: process, outcome and affect in human-computer interactions. *Canadian Journal of Library and Information Science*, 26 (4), 57-69. [On-line]. Available: http://www.ischool.utexas.edu/~adillon/publications/beyond_usability.pdf
- Draper, S. W. (1999). Analysing fun as a candidate software requirement. *Personal Technologies*, 3, 117-122.
- Dubs, R. (1993). Stehen wir vor einem Paradigmawechsel beim Lehren und Lernen? *Zeitschrift für Berufs- und Wirtschaftspädagogik* 5, 449-454.
- Dzida, W., Hoffmann, B., Freitag, R., Redtenbacher, W., Baggen, R., Geis, T., Beimel, J., Hurheiden, C., Hampe-Neteler, W., Hatwig, R., Peters, H. (2000). *Gebrauchstauglichkeit von Software. ErgoNorm: ein Verfahren zur Konformitätsprüfung von Software auf der Grundlage von DIN EN ISO 9241 Teile 10 und 11*. Schriftreihe der Bundeszentrale für Arbeitsschutz und Arbeitsmedizin. Forschung F1693, Dortmund.

- Eilks, I., Krille, B. & Flintjer, B. (2004). *Computer und Multimedia im Chemieunterricht heute - Eine Einordnung aus didaktischer und lerntheoretischer Sicht*. Broschüre der Gesellschaft Deutscher Chemiker, Fachgruppe Chemieunterricht, Arbeitsgruppe Computer im Chemieunterricht. [On-line]. Available: http://www.uni-frankfurt.de/didachem/gdchfg/cum_ck_2004.pdf
- Euler, D. (1994). (Multi)mediales Lernen - Theoretische Fundierungen und Forschungsstand. *Unterrichtswissenschaft*, 22, 291-311.
- Ferguson, D. L. (1992). Computers in Teaching and Learning: An Interpretation of Current Practices and Suggestions for Future Directions. In E. Scanlon & T. O'Shea (Eds.), *New Directions in Educational Technology* (pp. 34-50). Berlin: Springer.
- Ferguson, N. H. & Chapman, S. R. (1993). Computer-assisted instruction for introductory genetics. *Journal of Natural Resources and Life Sciences Education*, 22, 145-152.
- Fjeld, M., Juchli, B. M. & Vögli, B. (2003). Chemistry Education: A Tangible Interaction Approach. *Proceedings of INTERACT*, 287-294.
- Freudenreich, M. & Schulte, C. (2002). Von der Evaluation von Lernsoftware zur Gestaltung von Unterricht. *Online-Zeitschrift Medienpädagogik*. [On-line]. Available: http://www.medienpaed.com/02-1/freudenreich_schulte1.pdf
- Frey, K. (2002). *Allgemeine Didaktik*. Unveröffentlichtes Skript, Arbeitsunterlagen zur Vorlesung, 15. Auflage. Zürich.
- Fricke, R. (1995). Evaluation von Multimedia. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 401-413). Weinheim: Psychologie-Verl.-Union.
- Glowalla, U. & Häfele, G. (1995). Einsatz elektronischer Medien: Befunde, Probleme und Perspektiven. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 415-434). Weinheim: Psychologie-Verl.-Union.
- Glowalla, U. & Schoop, E. (1992). Entwicklung und Evaluation computerunterstützter Lehrsysteme. In U. Glowalla & E. Schoop (Hrsg.), *Hypertext und Multimedia. Neue Wege in der computerunterstützten Aus- und Weiterbildung* (S. 21-36). Heidelberg: Springer.
- Grizzle, J. E. (1965). The two-period change-over design and its use in clinical trials. *Biometrics*, 21, 469-480.
- Grund, S., Windlinger L. & Grote G. (2002). Wunsch und Wirklichkeit in der Nutzung eines webbasierten Kurses an der Universität. In M. Herczeg, H. Oberquelle & W. Prinz (Hrsg.), *Mensch & Computer 2002. Vom interaktiven Werkzeug zu kooperativen Arbeits- und Lernwelten*. Wiesbaden: Teubner-Verlag.
- Grüner, H. (1993). Evaluation und Evaluationsforschung im Bildungswesen. *Pädagogische Rundschau*, 47(1), 29-52.

-
- Hacker, W. & Richter, P. (1986). *Psychologische Bewertung von Arbeitsgestaltungsmassnahmen*. Berlin: Springer.
- Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshakti (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North Holland.
- Hasebrook, J. P. (1995). Lernen mit Multimedia. *Zeitschrift für Pädagogische Psychologie*, 9(2), 95-103.
- Hassenzahl, M., Beu, A. & Burmester, M. (2001). Engineering Joy. *IEEE Software*, 1&2, 70-76.
- Hegner, M. (2003). *Methoden zur Evaluation von Software*. IZ-Arbeitsbericht Nr. 29. Bonn: Informationszentrum Sozialwissenschaften. [On-line]. Available: http://www.gesis.org/Publikationen/Berichte/IZ_Arbeitsberichte/pdf/ab_29.pdf
- Huk, T. (2003). *Multimediales Lernen - Ein Überblick über die Forschungslandschaft*. Forschungsbericht (Nr. 51) aus dem Institut für Sozialwissenschaften (ISW), S. 1-34. TU Braunschweig.
- Hussy, W. & Jain, A. (2002). *Experimentelle Hypothesenprüfung in der Psychologie*. Göttingen: Hogrefe.
- Igbaria, M., Schiffman, S. J. & Wieckowski, T. J. (1994). The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology. *Behaviour & Information Technology*, 13, 349-361.
- ISO 9241-11 (1998). Ergonomic requirements for office work with visual display terminals - Part 11 Guidance on usability.
- Issing, L. J. (1995). Instruktionsdesign für Multimedia. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 195-220). Weinheim: Psychologie-Verl.-Union.
- Joliceur, K. & Berger, D. E. (1986). Do we really know what makes educational software effective? A call for empirical research. *Educational Technology*, 26(12), 7-11.
- Jonassen, D. H. (1991). What are cognitive tools? In P. A. M. Kommers et al. (Eds.), *Cognitive tools for learning*. Berlin: Springer.
- Jones, B. & Kenward, M.G. (2003). *Design and Analysis of Cross-Over Trials* (2nd edition). London: Chapman & Hall.
- Kerres, M. & de Witt, C. (2002). Quo vadis Mediendidaktik? Zur theoretischen Fundierung von Mediendidaktik. *Online-Zeitschrift Medienpädagogik* (2). [On-line]. Available: http://www.medienpaed.com/02-2/kerres_dewitt1.pdf

- Kerres, M. (2000). Information und Kommunikation beim mediengestützten Lernen. Entwicklungslinien und Perspektiven mediendidaktischer Forschung. *Zeitschrift für Erziehungswissenschaft*, 3 (1). 111-130.
- Kerres, M. (2001). *Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung* (2. Auflage). München: Oldenbourg.
- Kerres, M. (2004). Gestaltungsorientierte Mediendidaktik und ihr Verhältnis zur Allgemeinen Didaktik. In B. Dieckmann & P. Stadtfeld (Hrsg), *Allgemeine Didaktik im Wandel*. Bad Heilbrunn: Klinhardt Verlag.
- Kirakowski, J. & Corbett, M. (1993). SUMI: the Software Usability Measurement Inventory. *Britisch Journal of Educational Technology* 24(3), 210-212.
- Kirakowski, J. (1997) *Background notes on the SUMI questionnaire*. [Online]. Available: <http://sumi.ucc.ie/>
- Klimsa, P. (1995). Multimedia aus psychologischer und didaktischer Sicht. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 7-24). Weinheim: Psychologie-Verl.-Union.
- Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Education technology research and development*, 42 (2), S. 7-19.
- Kulik, C. C. & Kulik, J. A. (1991). Effectiveness of Computer-Based Instruction: An Updated Analysis. *Computers in Human Behavior*, 7, 75-94.
- Kulik, J. A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker & H. F. O'Neil Jr. (Eds.), *Technology assessment in education and training* (S. 9-34). Hillsdale: Erlbaum.
- Lang, P. (1985). *The cognitive Psychology of emotion: anxiety and the anxiety disorders*. Hillsdale: Erlbaum.
- Lehmacher, W. (1987). *Verlaufskurven und Cross-Over*. Berlin: Springer.
- Mader, G. & Stöckl, W. (1999). *Virtuelles Lernen. Begriffsbestimmung und aktuelle empirische Befunde*. Innsbruck: Studien Verlag.
- Maturana, H. R. & Varele, F. J. (1987). *Der Baum der Erkenntnis. Die biologischen Wurzeln des menschlichen Erkennens*. Bern: Scherz.
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge: MIT Press.
- Morell, D. (1992). The effects of computer assisted instruction on student achievement in high school biology. *School Science and Mathematics*, 92, 222-230.

- Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana: University of Illinois.
- Pantelidis, V.S. (1995). *Reasons to Use Virtual Reality in Education, VR in the Schools*. [Online]. Available: <http://www.soe.edu/vr/reas.html>
- Perelman, L. J. (1992). *School's out. A Radical New Formula for the Revitalization of America's Educational System*. New York: Avon Books.
- Pfendler, C. (1990). Zur Messung der mentalen Beanspruchung mit dem NASA-Task Load Index. *Zeitschrift für Arbeitswissenschaft*, 44, 158-163.
- Porteous, M., Kirakowski, J., Corbett, M. (1993). *SUMI user handbook*. Human Factors Research Group, University College Cork, Ireland.
- Reed, J.F. (2003). Crossover Designs in Lower Extremity Wounds. *Lower Extremity Wounds* 2(3), 158-163.
- Reiners, C. S., & Saborowski, J. (2001). Virtuelle Welten im Chemieunterricht. *Praxis der Naturwissenschaften - Chemie in der Schule*, 50, Heft 7, 9-12.
- Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C. & Barnes, C. (1999). Learning and building together in an immersive virtual world. *Presence*, 8(3), 247-263.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Russell, T. L. (1999). *The No Significant Difference Phenomenon*. Chapel Hill, NC: Office of Instructional Telecommunications, North Carolina State University.
- Salzman, M., Dede, C., Loftin, R. & Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual learning. *Presence*, 8(3), 293-316.
- Schallberger, U., Pfister, R. & Venetz, M. (1999). *Theoretische Rahmenüberlegungen zum Erlebens-Stichproben-Fragebogen (ESF) und zu den Operationalisierungen*. Arbeitsberichte aus dem Projekt "Qualität des Erlebens in Arbeit und Freizeit", Nr. 1. Zürich: Abteilung Angewandte Psychologie des Psychologischen Instituts der Universität.
- Schanda, F. (1995). *Computer-Lernprogramme: wie damit gelernt wird, wie sie entwickelt werden, was sie im Unternehmen leisten*. Weinheim: Beltz.
- Schaumburg H. & Rittmann, S. (2000). *Evaluation des Web-basierten Lernens - Ein Überblick über Werkzeuge und Methoden*. [On-line]. Available: <http://www.cmr.fu-berlin.de/~heike/papers/Eval-im-Internet.pdf>
- Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Zs., Encarnação, L. M., Gervautz, M. & Purgathofer, W. (2002). The Studierstube Augmented Reality Project. *PRESENCE - Teleoperators and Virtual Environments*, 11(1), 32-54.

- Schulmeister, R. (1997). *Grundlagen hypermedialer Lernsysteme. Theorie - Didaktik - Design* (2. Auflage). München: Oldenbourg.
- Schulmeister, R., Vollmers, B., Gücker, R. & Nuyken, K. (2003). *Konzeption und Durchführung der Evaluation einer virtuellen Lernumgebung: Das Projekt Methodenlehre-Baukasten*. [On-line]. Available: http://www.uni-kassel.de/fb1/mediafb1/dgfemedien/Seiten/Texte/tagungen/herbsttagung03/01_Schulmeister/Magdeburg_Vortrag_301003_2e.pdf.
- Schwan, S. & Buder, J. (2002). Lernen und Wissenserwerb in virtuellen Realitäten. In G. Bente, N. C. Krämer & A. Petersen (Hrsg.), *Virtuelle Realitäten* (S. 109-132). Göttingen: Hogrefe.
- Seel, N. M. (2000). *Psychologie des Lernens*. München: Ernst Reinhardt.
- Shneiderman, B. (1997). *Direct Manipulation for Comprehensible, Predictable, and Controllable User Interfaces*. In Proceedings of IUI97, International Conference on Intelligent User Interfaces (pp. 33-39), Orlando (FL).
- Shuell, T.J. (1987). Cognitive conceptions of learning. *Review of educational research*, 56, 411-436.
- Stangl, W. (2004). *Lernen mit dem Computer*. In [werner-stangl]s arbeitsblätter. [On-line]. Available: <http://arbeitsblaetter.stangl-taller.at/LERNEN/Computerlernen.shtml>
- Strittmatter, P. & Mael, D. (1995). Einzelmedium, Medienverbund und Multimedia. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 47-61). Weinheim: Psychologie-Verl.-Union.
- Sumfleth, E. & Kummer, T. (2001). Lernen mit Hypertexten zur Einführung in einen Themenbereich - Beispiel Seife. *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 147-165.
- Tabachnik, B. G. & Fidell, L. S. (1996). *Using Multivariate Statistics* (3rd Edition). New York: HarperCollins.
- Tergan, S.-O. (1997). Lernen mit Texten, Hypertexten und Hypermedien. Retrospektive und State of the Art. In H. Gruber & A. Renkl (Hrsg.), *Wege zum Können. Determinanten des Kompetenzerwerbs* (S. 236-249). Bern: Huber.
- Toutenburg, H. (1994). *Versuchsplanung und Modellwahl*. Heidelberg: Physica-Verlag.
- Trautmann, T. (2004). *Der regulatorische Fokus des Nutzers: Motivationale Determinanten bei der Bewertung von Software*. Diplomarbeit. [On-line]. Available: <http://www.tibix.de/diplomarbeit/abstract.html>
- Tulodziecki, G., Hagemann, W., Herzig, B., Leufen, S. & Mütze, C. (1996). *Neue Medien in den Schulen: Projekte-Konzepte-Kompetenzen*. Gütersloh: Bertelsmann Stiftung.
- Ulich, E. (1994). *Arbeitspsychologie* (3. Auflage). Stuttgart: Schäfer-Poeschel.

- Unema, P., Rötting, M., Sepher-Willeberg, M., Strümpfel, U. & Kopp, U. (1988). *Der NASA Task Load Index: Erste Ergebnisse mit der deutschen Fassung*. In Gesellschaft für Arbeitswissenschaft e.V. (Hrsg.). Jahresdokumentation 1988 der Gesellschaft für Arbeitswissenschaft e.V. - Bericht zum 34. Arbeitswissenschaftlichen Kongress an der RWTH Aachen. Köln: O. Schmidt.
- Vögtli, B. (2002). *Augmented Collaboration*. Diploma thesis of Benedikt Vögtli, March 2002, HyperWerk FHBB. Diplomarbeit.
- Wainwright, C. L. (1989). The effectiveness of a computer-assisted instruction package in high school chemistry. *Journal of Research in Science Teaching*, 26, 275-290.
- Will, H., Winteler, A. & Krapp, A. (1986). Von der Erfolgskontrolle zur Evaluation. In H. Will, A. Winteler & A. Krapp (Hrsg.), *Evaluation in der beruflichen Aus- und Weiterbildung. Konzepte und Strategien* (S. 11-42). Heidelberg: Sauer.
- Winn, W.A. (1993). *Conceptual Basis for Educational Applications of Virtual Reality*. Technical Report TR 93-9. [On-line]. Available: <http://www.hitl.washington.edu/publications/r-93-9/>
- Wottawa, H. & Thierau, H. (1990). *Lehrbuch Evaluation*. Bern: Huber.
- Yalcinalp, S., Geban, O. & Ozkan, I. (1995). Effectiveness of using computer-assisted supplementary instruction for teaching the mole concept. *Journal of Research in Science Teaching*, 32, 1083-1095.
- Zwisler, R. (1998). *Virtuelle Realität und die Rolle von Haptik*. Institut für Psychologie, Universität Regensburg. [On-line]. Available: <http://www.zwisler.de/scripts/haptics/haptics.html>

Anhang

Anhang A: Aufgabenblätter der beiden Versuche

5. Stellen Sie nun längere Moleküle mit 4, 5 resp. 6 C-Atomen dar und notieren Sie jedes Mal die Summen- sowie die Strukturformel:

C-Kette:	Summenformel:	Strukturformel:
4 C:		
5 C:		
6 C:		

6. Stellen Sie nun ein Molekül mit insgesamt vier C-Atomen her worin eine Verzweigung vorkommt. Die C-Kette nennt man das „Gerüst“. Wenn Sie verschiedene Moleküle mit derselben Summenformel aber verschiedenen Strukturen hergestellt haben, haben Sie verschiedene „Isomere“ dargestellt.

Vergleichen Sie die Summenformel mit dem andern C₄-Molekül.
Was stellen Sie fest?

Wie viele solche verzweigte C₄-Moleküle sind denkbar?

7. Lösen Sie dieselbe Aufgabe für das C₅-Molekül:

Vergleichen Sie die Summenformel mit dem andern C₅-Molekül.
Was stellen Sie fest?

Wie viele solche verzweigte C₅-Moleküle sind denkbar?

8. Stellen Sie eine allgemeine Regel auf über die Zahl der C-Atome im Molekül und die Zahl der möglichen Isomere.

9. Definieren Sie „Isomere“:

- Lernaufgaben für den ersten Versuch -

Name:

Arbeitsauftrag: Grundlagen der Organischen Chemie 1

Ziel: Sie können

- die Prinzipien nach denen die organischen Moleküle aufgebaut sind beschreiben.
- zwischen Struktur- und Summenformel unterscheiden.
- „Isomere“ definieren.

1. Konstruieren Sie das einfachste Molekül das 1 C enthält und dessen Bindungen alle mit H besetzt sind.

Wie heisst dieser Stoff?	
Zeichnen Sie eine Strukturformel des Moleküls:	
Welche Summenformel hat das Molekül?	
Aggregatzustand bei Raumtemperatur:	
Löslich in:	
Verwendung:	
Vorkommen:	

2. Stellen Sie den entsprechenden Stoff mit 2 C-Atomen dar.

Wie heisst dieser Stoff?	
Zeichnen Sie eine Strukturformel des Moleküls:	
Welche Summenformel hat das Molekül?	
Aggregatzustand bei Raumtemperatur:	
Löslich in:	
Verwendung:	
Vorkommen:	

3. Stellen Sie den entsprechenden Stoff mit 3 C-Atomen dar.

Wie heisst dieser Stoff?	
Zeichnen Sie eine Strukturformel des Moleküls:	
Welche Summenformel hat das Molekül?	
Aggregatzustand bei Raumtemperatur:	
Löslich in:	
Verwendung:	
Vorkommen:	

4. Notieren Sie noch einmal die Namen dieser ersten drei Moleküle, Sie müssen diese auswendig können.

--	--

5. Verlängern Sie die C-Kette des vorherigen Moleküls um ein C und ergänzen Sie es wieder mit H-Atomen. (Eventuell ist es am einfachsten, ganz neu zu beginnen.)

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff (2 Namen)?
Was für einen Gefrierpunkt hat er?
Nennen Sie gefährliche Eigenschaften.

6. Beginnen Sie mit einem ganz neuen Molekül: An ein C-Atom fügen Sie ein N-Atom an (Einfachbindung). Ergänzen Sie alle noch freien Bindungen mit H-Atomen.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff?
Was für einen Aggregatzustand hat er?
Nennen Sie gefährliche Eigenschaften.

7. Stellen Sie eine Kette aus drei C her, fügen Sie an das mittlere C ein Sauerstoffatom mittels einer Doppelbindung an. Ergänzen Sie das Molekül mit H-Atomen.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff (2 Namen)?
Was für einen Aggregatzustand hat er?
Nennen Sie gefährliche Eigenschaften.

8. Wieder ein ganz neues Molekül: Eine Kette aus zweimal zwei C, unterbrochen von einem Sauerstoff in der Kette. Auch hier wieder: Alle freien Bindungen mit H ergänzen.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff?
Was für einen Aggregatzustand hat er?
Nennen Sie gefährliche Eigenschaften.

2

- Lernaufgaben für den zweiten Versuch -

Name:

Arbeitsauftrag: Grundlagen der Organische Chemie 2
Ziel: Nach dieser Übung sollen Sie einige Vertreter verschiedener Stoffgruppe erkennen und aufzeichnen können.

1. Stellen Sie ein Molekül mit der Formel $H_2C=CH_2$ (Doppelbindung!) her.

Wie nennt man diesen Stoff?
Was für einen Aggregatzustand hat er?

2. Beginnen Sie wieder mit einem neuen Molekül (Löschkarte): Stellen Sie ein Methanmodell her (CH_4) und ersetzen Sie ein H durch eine $-OH$ -Gruppe.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff?
Was für einen Aggregatzustand hat er?
Welche Gefahren sind mit diesem Stoff verbunden?

3. Ändern Sie das Modell aus Aufgabe 2 so ab, dass der Sauerstoff mit beiden Bindungen am C befestigt ist.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff (2 Namen)?
Was für einen Aggregatzustand hat er?
Nennen Sie gefährliche Eigenschaften.

4. Nun ergänzen Sie das Molekül: Es soll jetzt neben dem doppelt gebundenen Sauerstoff noch eine $-OH$ -Gruppe tragen. Wieder haben wir einen Vertreter einer neuen Stoffgruppe.

Zeichnen Sie die Strukturformel:
Wie nennt man diesen Stoff?
Was für einen Aggregatzustand hat er?
Nennen Sie gefährliche Eigenschaften.

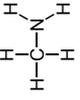
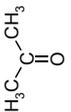
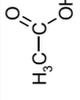
1

Anhang B: Kurzttests der beiden Versuchen

- Kurztest für den ersten Versuch -

Name:	
Kurztest: Grundlagen der Organischen Chemie 1	
Zeit: 10 Minuten	
1.	Wie heisst dieser Stoff: CH ₄ ?
2.	Zeichnen Sie die Strukturformel von CH ₄ .
3.	Welchen Aggregatzustand hat CH ₄ bei Raumtemperatur?
4.	Zeichnen Sie die Strukturformel des Ethans.
5.	Geben Sie die Summenformel des Ethans an.
6.	Zeichnen Sie Propan.
7.	Geben Sie die Summenformel des Propan an.
8.	Tragen Sie Aussagen über verschiedene Isomere desselben Stoffes ein: Ihre Summenformel: Ihre Molmasse: Ihre Strukturformel:
9.	Zeichnen Sie nach denselben Prinzipien wie bisher drei Isomere des Moleküls mit 6 C-Atomen („Hexan“).
10.	Welche Summenformel hat Hexan?

- Kurztest für den zweiten Versuch -

Name:	
Kurztest: Grundlagen der Organischen Chemie 2	
Zeit: 10 Minuten	
1.	Wie nennt man diesen Stoff: H ₂ C=CH ₂ ? Was für einen Aggregatzustand hat er?
2.	Zeichnen Sie ein Formaldehydmolekül Welchen anderen Namen hat Formaldehyd auch noch?
3.	Zeichnen Sie ein Ameisensäuremolekül. Nennen Sie eine typische, gefährliche Eigenschaft der Ameisensäure.
4.	Wie heisst dieser Stoff? 
5.	Was für einen Aggregatzustand hat er? Wie heisst dieser Stoff (2 Namen)? 
6.	Wozu verwendet man diesen Stoff? Nennen Sie eine gefährliche Eigenschaft dieses Stoffes. Zeichnen Sie Diethylether. Nennen Sie eine typische, gefährliche Eigenschaft von Diethylether.
7.	Wie heisst dieser Stoff (2 Namen)? 
8.	Wo liegt sein Gefrierpunkt (ankreuzen)? - 12,3°C 2,1°C 16,5°C 24,9°C Zeichnen Sie Methanol. Nennen Sie eine gefährliche Eigenschaft von Methanol.

Anhang C: Nasa Task Load Index

Bemerkung: Abgedruckt sind im Folgenden (Anhang C bis Anhang F) die Fragebögen zur Beurteilung der Augmented Chemistry - Lernsituation. Natürlich bestanden diese Fragebögen gleichermassen für das Kugel-Stab-Modell.

Beschreibung der sechs Beanspruchungsdimensionen

Bitte machen Sie sich mit den folgenden sechs Beanspruchungsdimensionen vertraut. Lesen Sie dazu unten die Erläuterungen sorgfältig durch. Es ist wichtig, dass Sie die Dimensionen verstanden haben. Nachher werden Sie das Arbeiten mit Augmented Chemistry (AuC) anhand dieser sechs Dimensionen bewerten. Dieses Blatt können Sie jederzeit zur Hilfe nehmen.

AuC = Augmented Chemistry

Geistige Beanspruchung	Wie sehr waren Sie durch geistige Arbeit und Wahrnehmungsleistungen beansprucht (z.B. Denken, Entscheiden, Erinnern, Hinsehen, Suchen, etc.)? War das Arbeiten mit AuC in dieser Hinsicht eher leicht oder anspruchsvoll, einfach oder komplex, erforderte es hohe Genauigkeit oder war es fehlerolerant?
Körperliche Beanspruchung	Wie sehr waren Sie durch körperliche Arbeit beansprucht (z.B. Drücken, Ziehen, Drehen, Steuern, Aktivieren, etc.)? War das Arbeiten mit AuC in dieser Hinsicht eher leicht oder anspruchsvoll, einfach oder anstrengend, locker oder energisch, erholend oder mühselig?
Zeitliche Beanspruchung	Wie sehr haben Sie sich unter Zeitdruck gefühlt beim Lösen der Aufgaben mit AuC? War das Arbeitstempo eher langsam und gemächlich oder schnell und hektisch?
Anstrengung	Wie hart mussten Sie arbeiten (geistig und körperlich), oder wie sehr mussten Sie sich anstrengen, um Ihre Leistung zu erbringen?
Aufgabenbewältigung	Wie erfolgreich konnten Sie die vorgegebenen Aufgaben mit AuC lösen? Wie zufrieden waren Sie dabei mit Ihrer Leistung?
Frustration	Wie unsicher, entmutigt, genervt, gestresst und verärgert (im Gegensatz zu sicher, bestätigt, zufrieden und entspannt) fühlten Sie sich beim Arbeiten mit AuC?

(AuC)

- 2 -

Arbeitsbelastung beim Lernen mit Augmented Chemistry
(NASA Task Load Index)

Mit diesem Fragebogen wird das Ausmass der „Arbeitsbelastung“ untersucht, die Sie bei der Aufgabenbearbeitung mit Augmented Chemistry empfunden haben. Anhand von insgesamt sechs Beanspruchungsdimensionen sollen Sie Ihre Erfahrung im Umgang mit Augmented Chemistry beurteilen.

Die sechs Beanspruchungsdimensionen sind:

- Geistige Beanspruchung
- Körperliche Beanspruchung
- Zeitliche Beanspruchung
- Anstrengung
- Aufgabenbewältigung
- Frustration

Aufbau des Fragebogens:

- **Teil A:** Beanspruchungshöhe: Hier bewerten Sie, wie hoch die Beanspruchung in den einzelnen Dimensionen war.
- **Teil B:** Beanspruchungsstruktur: Hier bewerten Sie die relative Wichtigkeit der sechs Dimensionen für die empfundene Gesamtbeanspruchung

Lesen Sie nun auf folgender Seite die Beschreibung der sechs Beanspruchungsdimensionen durch und füllen Sie anschliessend den Fragebogen aus.

(AuC)

- 1 -

Name: Datum:

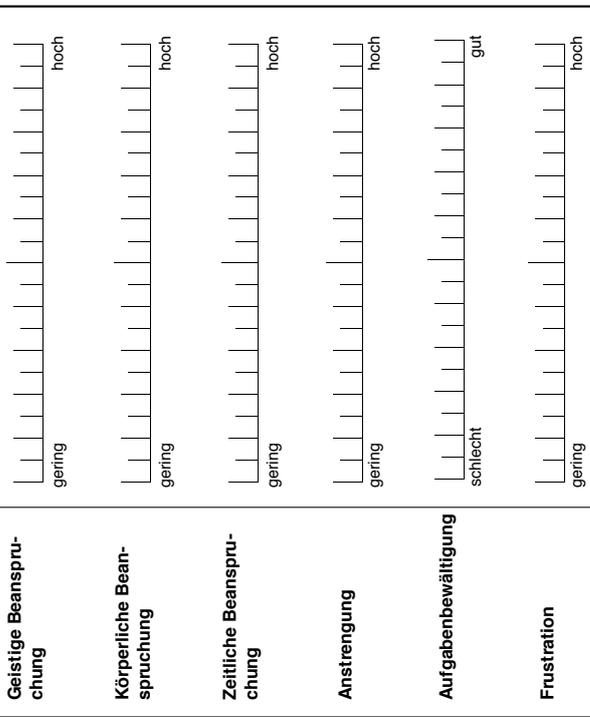
Teil A: Beanspruchungshöhe

Wie haben Sie das Arbeiten mit Augmented Chemistry (AuC) erlebt? Wie hoch war dabei die Beanspruchung in den einzelnen Dimensionen? Markieren Sie dazu auf den folgenden Skalen, in welchem Masse Sie sich in den sechs genannten Dimensionen beansprucht oder gefordert gesehen haben:



Diese Person fühlte sich durch das Arbeiten mit AuC eher wenig geistig beansprucht.

Wie schätzen Sie Ihre Belastung durch AuC in folgenden Dimensionen ein?



- 3 - (AuC)

Name: Datum:

Teil B: Beanspruchungsstruktur

Welche relative Bedeutung für die empfundene Gesamtbeanspruchung hatten die sechs Beanspruchungsdimensionen für Sie?

Im Folgenden werden jeweils zwei der sechs Beanspruchungsdimensionen in allen möglichen Kombinationen gegenübergestellt. Geben Sie bitte jeweils an, welche der beiden Beanspruchungsdimensionen bedeutsamer war für Ihre empfundene Gesamtbeanspruchung bei der Aufgabenbearbeitung mit Augmented Chemistry (AuC).

Es geht zunächst nicht darum, wie hoch die Beanspruchung in den einzelnen Dimensionen war, sondern **wie wichtig** die jeweilige Dimension für das Gesamttempfinden war!

Beispiel:

Geistige Beanspruchung Frustration

Für diese Person hat ihre Frustration mehr zur gesamthaften Beanspruchung beigetragen als ihre empfundene geistige Beanspruchung bei der Aufgabenbearbeitung. Der Frustration wird in diesem Fall mehr Gewicht beigemessen als der geistigen Beanspruchung.

Überlegen Sie sich nun für jedes unten präsentierte Paar:

Welche der beiden Dimensionen hat mehr zu Ihrer Gesamtbeanspruchung bei der Aufgabenbearbeitung mit AuC beigetragen?

(Zur Erinnerung: Die Bedeutung der folgenden Dimensionen können Sie jederzeit auf Seite 2 nachlesen.)

Körperliche Beanspruchung Zeitliche Beanspruchung

Anstrengung Geistige Beanspruchung

Frustration Körperliche Beanspruchung

Anstrengung Frustration

(AuC)

- 4 -

Name: Datum:

Geistige Beanspruchung	<input type="checkbox"/>	Zeitliche Beanspruchung
Körperliche Beanspruchung	<input type="checkbox"/>	Anstrengung
Zeitliche Beanspruchung	<input type="checkbox"/>	Aufgabenbewältigung
Frustration	<input type="checkbox"/>	Geistige Beanspruchung
Zeitliche Beanspruchung	<input type="checkbox"/>	Frustration
Aufgabenbewältigung	<input type="checkbox"/>	Anstrengung
Anstrengung	<input type="checkbox"/>	Zeitliche Beanspruchung
Frustration	<input type="checkbox"/>	Aufgabenbewältigung
Aufgabenbewältigung	<input type="checkbox"/>	Körperliche Beanspruchung
Geistige Beanspruchung	<input type="checkbox"/>	Aufgabenbewältigung
Geistige Beanspruchung	<input type="checkbox"/>	Körperliche Beanspruchung

Vielen Dank fürs Ausfüllen!

- 5 - (AUC)

Anhang D: Self-Assessment-Manikin (SAM)

Name:

3. Juni 2004 (AuC)

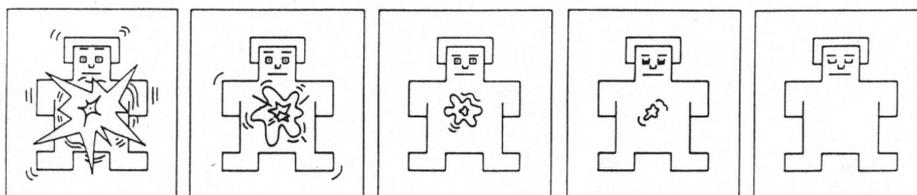
Fragebogen zur Befindlichkeit

(SAM)

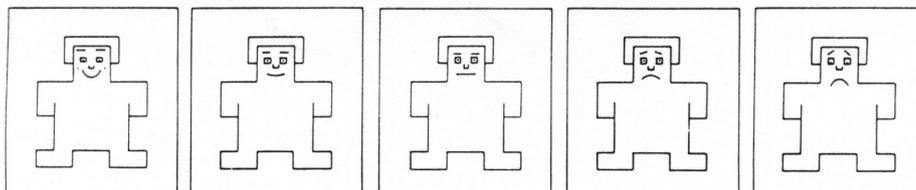
Wie fühlen Sie sich nach dem Arbeiten mit Augmented Chemistry?

Bitte markieren Sie unten jeweils die entsprechende Figur mit einem Kreuz. Ein Ankreuzen zwischen zwei Figuren ist erlaubt.

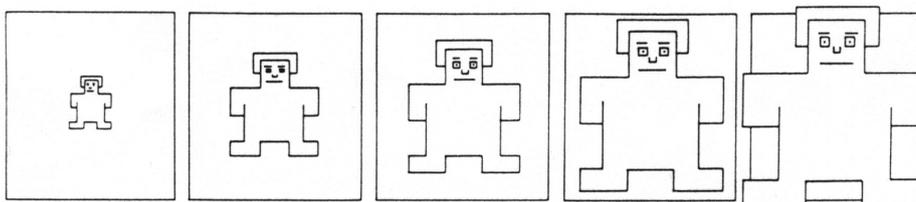
1. Wie stark sind Sie emotional erregt?



2. Wie angenehm bzw. unangenehm ist Ihr Gefühl?



3. Wie sehr haben Sie die Situation unter Kontrolle?



Name: Datum:

C

Was hat Ihnen beim Arbeiten mit Augmented Chemistry besonders gut gefallen, was hat Ihnen eher Mühe bereitet?

Folgendes hat mir besonders gut gefallen:

.....
.....
.....
.....
.....

Folgendes hat mir nicht so gut gefallen:

.....
.....
.....
.....
.....

Wenn Sie Augmented Chemistry eine Gesamtnote von 1 (ungenügend) bis 6 (sehr gut) geben könnten, wie würde diese lauten?

Note:

Falls Sie Ideen zur Erweiterung oder Verbesserung von Augmented Chemistry haben, möchte ich Sie bitten dies auf der Rückseite zu notieren.

Vielen Dank für Ihre Unterstützung!

- 3 -

(AuC)

Anhang F: Fragebogen zur Präferenz eines Lernsystems

Name:

Fragebogen zur Präferenz eines Lernsystems

Dieser Fragebogen erfasst, ob Sie unter gewissen Gesichtspunkten lieber mit Augmented Chemistry oder lieber mit dem Kugel-Stab-Modell gearbeitet haben. Denken Sie bitte zurück an die Aufgabenbearbeitung mit Augmented Chemistry und dem Kugel-Stab-Modell. Vergleichen Sie Ihr Arbeiten mit den beiden Lernsystemen und beantworten Sie dann folgende Fragen.

Markieren Sie bitte mit einem Kreuz Ihr jeweils bevorzugtes Lernsystem. Falls Sie unentschieden sind oder keine Präferenz haben, kreuzen Sie bitte das Kästchen ganz rechts an.

		unent- schieden
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... hat mir mehr Spass bereitet.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... fand ich leichter zu erlernen.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... fand ich einfacher zu bedienen.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... fand ich angenehmer im Umgang.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... war für mich die bessere Unterstützung beim Lösen der Aufgaben.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... konnte mir den Lerninhalt besser veranschaulichen und verständlich machen.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... bot mir den besseren Zugang zum Lerninhalt.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... würde ich in Zukunft lieber anwenden.	<input type="checkbox"/>
<input type="checkbox"/> Augmented Chemistry... <input type="checkbox"/> Kugel-Stab-Modell...	... finde ich effizienter zum Lernen.	<input type="checkbox"/>

Für weitere Anregungen oder Bemerkungen verwenden Sie bitte die Rückseite!

Vielen Dank fürs Ausfüllen!

Anhang G: Software Usability Measurement Inventory (SUMI)

Name: Datum:

	 nicht einverstanden	 unentschieden	 einverstanden
1. Die Software reagiert zu langsam auf die Eingaben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Ich würde die Software meinen Kollegen empfehlen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Die Anweisungen und Antworten (Bedienungshinweise / Benutzerführung) sind hilfreich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Die Software hängt sich gelegentlich unerwartet auf.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Der erste Umgang mit dieser Software ist voller Probleme.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Manchmal weiss ich nicht, was ich als nächstes mit der Software machen kann.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Die Arbeit mit dieser Software macht mir Spass.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Ich finde, dass die angebotenen Help-Informationen (Bedienungsanleitung) nicht sehr hilfreich sind.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Ein Neustart ist schwierig, falls sich das System aufhängt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Es dauert zu lange, um die Software-Kommandos zu erlernen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Manchmal frage ich mich, ob ich die richtigen Kommandos verwende.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Es ist befriedigend, mit dieser Software zu arbeiten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Das System präsentiert die Informationen klar und verständlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Ich fühle mich sicherer, wenn ich nur wenige vertraute Befehle und Operationen verwende.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Die Dokumentation (Bedienungsanleitung) ist sehr informativ.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Die Software unterbricht mich in meinem normalen Arbeitsfluss.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Die Arbeit mit dieser Software stimuliert mich geistig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

© 2004, Human Factors Research Group, Ireland. Student licence for Mrs. K. Bötschi, till 12/2004

- 2 -

Name: Datum:

Die Benutzbarkeit von Augmented Chemistry

(SUMI: Software Usability Measurement Inventory)

Allgemeine Information:
Dieser Fragebogen wird nur von denjenigen ausgefüllt, die mit Augmented Chemistry gearbeitet haben.

Zum Fragebogen:

- Der Fragebogen enthält 50 Aussagen. Hinter jeder Aussage finden Sie drei Kästchen.
- Kreuzen Sie das linke Kästchen (☹) an, wenn Sie mit der Aussage grundsätzlich nicht einverstanden sind.
- Kreuzen Sie das mittlere Kästchen (☺) an, wenn Sie sich nicht entscheiden können, oder wenn die Aussage für Ihre Software und Ihre Situation gegenstandslos ist.
- Kreuzen Sie das rechte Kästchen (😊) an, wenn Sie mit der Aussage grundsätzlich einverstanden sind.

Es handelt sich hierbei um einen standardisierten Fragebogen, der für verschiedene Software gleichermassen angewendet werden kann. Deshalb kommt es vor, dass gewisse Aussagen nicht auf Augmented Chemistry angewandt werden können. Lassen Sie sich dadurch nicht irritieren und kreuzen Sie bei solchen Aussagen „unentschieden“ (☺) an.

Wenn im Folgenden von „Software“ die Rede ist, wird immer auf Augmented Chemistry Bezug genommen.

© 2004, Human Factors Research Group, Ireland. Student licence for Mrs. K. Bötschi, till 12/2004

- 1 -

Name:	Datum:		
35. Es ist schwierig, die Anwendung neuer Funktionen zu erlernen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. Es bedarf zu vieler Einzelschritte, um eine Funktion auszuführen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37. Diese Software hat mir gelegentlich Kopfschmerzen bereitet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38. Die Warnungen zur Fehlervermeidung sind ungenügend.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39. Es ist leicht, mit der Software genau das zu machen, was ich möchte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
40. Ich werde niemals lernen, alle durch die Software angebotenen Optionen anzuwenden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41. Die Software hat nicht immer genau das getan, was ich erwartet habe.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42. Diese Software hat eine sehr attraktive Aufmachung.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
43. Der Umfang oder die Qualität der Help-Informationen variiert im System.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
44. Es ist relativ einfach, von einer Teilaufgabe zu einer anderen zu wechseln (z.B. von „Molekülbau“ zu „Bibliothek-Modus“).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45. Man vergisst sehr schnell, wie man Aufgaben mit dieser Software erledigt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
46. Die Software verhält sich gelegentlich unverständlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47. Diese Software ist wirklich sehr umständlich zu handhaben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48. Man sieht mit einem Blick, welche Optionen (Handlungsmöglichkeiten) im Augenblick zur Verfügung stehen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49. Die Eingabe und der Zugang zu Dateien im System ist nicht einfach.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
50. Ich brauche sehr oft Unterstützung, wenn ich diese Software benutze.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vielen Dank fürs Ausfüllen!

Name:	Datum:		
18. Es ist nie genug Information auf dem Bildschirm, wenn sie gebraucht wird.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Ich habe das Gefühl, die Software zu beherrschen, wenn ich sie benutze.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Ich ziehe es vor, bei den Möglichkeiten zu bleiben, die ich am besten kenne.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Ich denke, diese Software ist inkonsistent.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. Ich möchte diese Software nicht jeden Tag verwenden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. Ich verstehe die von der Software bereitgestellten Informationen und arbeite damit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Diese Software ist umhandlich, wenn ich etwas machen möchte, was nicht Standard ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Ich muss zu viel lesen, bevor ich die Software anwenden kann.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Aufgaben können mit dieser Software direkt und ohne Umschweife ausgeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Die Anwendung dieser Software ist frustrierend.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Diese Software hat mir geholfen, alle Probleme zu lösen, die ich beim Umgang damit hatte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Die Geschwindigkeit der Software ist hoch genug.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Ich muss regelmässig Einzelheiten in der Bedienungsanleitung nachschlagen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Die Bedürfnisse der Benutzer sind vollständig berücksichtigt worden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Manchmal war ich sehr angespannt während der Anwendung dieser Software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. Die Organisation der Menüs und Informationslisten erscheint sehr logisch.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
34. Die Software erlaubt dem Anwender sparsame Eingabebefehle.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>