

Beurteilungskriterien für Lernsoftware

von Andreas Holzinger, IMI, Uni Graz
Ein Service des Serverprojektes Joanneum
im Rahmen der Initiative

„Neue Medien an der Lehre für Universitäten und Fachhochschulen“

*„Evaluation tests the
usability and functionality
of an interactive system.“*

**Alan Dix et.al.
Lancaster University
(1998)**



*Alan Dix und
Andreas Holzinger vor
dem Grazer Uhrturm,
(HCI-Symposium,
2001)*

Grundlagen der Evaluation

Checkliste

Kriterien, Konzepte

Benutzer von Lernsoftware wollen rasch über die Qualität dieser Software bescheid wissen. Um Lernsoftware schnell, sicher und auch vergleichbar beurteilen zu können, sind einige elementare Grundlagen über Evaluation notwendig. Dieses Skript versucht, diese Grundlagen kurz und knapp zu vermitteln. Die Checkliste auf Seite 12 kann in der Praxis verwendet werden. Anhand einer **analytischen Vorgehensweise** wird diese Checkliste mit dem notwendigen Basiswissen „hinterlegt“. Dabei werden einige Konzepte zur Evaluierung von Lernsoftware vorgestellt. Die Literaturliste am Ende soll die Vertiefung ermöglichen.

Über den Autor:

Ing. Mag.Mag. Dr. Andreas Holzinger ist seit mehr als 20 Jahren im Bereich IT/Informatik tätig und arbeitet, lehrt und forscht derzeit am Institut für medizinische Informatik, Statistik und Dokumentation der Uni Graz im Fach Informationssysteme mit besonderer Berücksichtigung der Mensch-Maschine Kommunikation (Human-Computer Interaction, HCI).

Zur Ausbildung: Radio- und Fernsehtechniker (1981); Werkmeisterschule für Industrie-Elektronik mit Lehrlingsausbilderprüfung (1983); College of Further Education Bournemouth (UK) mit Schwerpunkt Computertechnik (1985/86); Ing. für Nachrichtentechnik (1990); Diplom in Erwachsenenbildung (Dip.Ed., 1991); Studien der Nachrichtentechnik, Physik und Psychologie (Mag.rer.nat., 1995) sowie Medienpädagogik und Soziologie (Mag.phil., 1996) an der TU und Uni Graz. Promotion mit „summa cum laude“ auf dem Gebiet der Kognitionswissenschaft (Dr.phil., 1997).

Er ist u.a Mitglied der ACM, IEEE, AACE, GI, DGP, GMW und Vorstandsmitglied der Austrian Computer Society (OCG). Dr. Holzinger ist österreichischer Experte in der EU in der eLearning – eEurope Initiative.

Homepage des Autors: <http://www-ang.kfunigraz.ac.at/~holzinge>

Homepage von Basiswissen Multimedia: www.basiswissen-multimedia.at

1 Grundlagen	6
1.1 Definition Evaluation	6
1.2 Arten von Evaluation	7
1.3 Mittel der Software-Evaluation	8
1.4 Skalen	9
1.5 Interpretation von Messwerten	9
2 Checklisten	10
2.1 Kriterienkataloge	10
2.2 Beurteilungsverfahren	11
2.3 Checkliste für Multimedia	12
3 Erläuterung der Checkliste	14
3.1 Allgemeine Beschreibung	14
3.1.1 A01 Zielgruppe	15
3.1.2 A02 Preis-Leistungs-Verhältnis	16
3.1.3 Technologischer Hintergrund	16
3.1.4 T01 Systemressourcen	17
3.1.5 T02 Ladezeiten	17
3.2 HCI-Aspekte	17
3.2.1 H01 Bildschirmaufteilung	18
3.2.2 H02 Screendesign	18
3.2.3 H03 Farben	19
3.2.4 H04 Helligkeit	20
3.2.5 H05 Formkonstanz	20
3.2.6 H06 Bedienerfreundlichkeit	20
3.2.7 H08 Textgröße	21
3.2.8 H09 Textlesbarkeit	21
3.2.9 H10 Symbolerkennung	22
3.3 Multimedialer Hintergrund	23
3.3.1 M01 Grafikqualität	23
3.3.2 M02 Klangqualität	23
3.3.3 M03 Sprachqualität	23
3.3.4 M04 Animationsqualität	23
3.3.5 M05 Filmqualität	23
3.3.6 M06 Interaktivität	24
3.4 Didaktik und Lernpsychologie	25
3.4.1 L01 Motivation	25
3.4.2 L02 Arousal	25
3.4.3 L03 Lerneffizienz	26
3.4.4 L04 Lernatmosphäre	26
3.4.5 L05 Lernintensität	27
3.4.6 L06 Neugierde	27
3.4.7 L07 Überraschung	27
3.4.8 L08 Humor	28
3.4.9 L09 Erwartungshaltung	28
3.4.10 L10 Kreativität	28
3.4.11 L11 Schwierigkeit	28
3.4.12 L12 Feedback	29
3.4.13 L13 Lernzielklarheit	29
3.4.14 L14 Lernzielerreichbarkeit	29
3.4.15 L15 Übungsmöglichkeit	29
3.4.16 L16 Abwechslung	30
3.4.17 Was ist noch zu beachten?	30
3.5 Kognitionspsychologischer Hintergrund	31
3.5.1 K01 Wissensvermittlung	31
3.5.2 K02 Inzidentielles Lernen	31
3.5.3 K03 Problemlösen	31
3.5.4 Gesamturteil	32
4 Zusammenfassung	33
5 Literatur	34
5.1 Bücher	34
5.2 Artikel	36
5.3 Books in English	36
5.4 Articles in English	37
5.5 Journals	38
5.6 Internet-Links	39

Entstehungsgeschichte dieses Skriptums. Als Vorwort.

Das „Lernen mit Computern“ begeisterte mich seit meinem Einstieg in die Informationstechnik im Jahre 1978. Ab 1995 beschäftigte ich mich auch wissenschaftlich mit „professioneller“ Lernsoftware – damals noch großteils auf CD-ROM. Besonders in der Medizin wurden multimediale Lernprogramme zunehmend eingesetzt. Dabei tauchte von den verschiedensten Leuten (Studierenden, Lehrenden, aber auch Schülern, Lehrern, Eltern, Freunden und Bekannten) immer wieder die gleiche Frage auf: „Wie kann ich Lernsoftware (möglichst schnell) hinsichtlich deren Qualität beurteilen?“ Da ich eine Hilfestellung zur (fachneutralen) Beantwortung dieser Frage geben wollte, entwickelte ich eine (möglichst einfache) Checkliste zur Beurteilung von Lernsoftware. Oberste Prämisse war dabei, dass die Checkliste auf einer DIN A4 Seite Platz hat. Aus dieser Checkliste entstand im Verlauf meiner Expertentätigkeit beim bm:bwk ein „Manual zur Beurteilung von Lernsoftware“ und das „Modul 6: Evaluation“ im Band 3 meiner Buchreihe „Basiswissen Multimedia“ (www.basiswissen-multimedia.at).

Da die Frage auch jetzt immer noch auftaucht, möchte ich eine überarbeitete Version dieser Checkliste für die Leserinnen und Leser des Newsletters „NML“ zur Verfügung stellen. Dabei betone ich aber, dass diese lediglich zur einfachen und raschen Beurteilung von einzelnen Lernprogrammen dient und nicht z.B. zur Beurteilung von Lernplattformen (das ist z.B. sehr ausführlich von meinem Kollegen PETER BAUMGARTNER in seinem „e-Learning Praxishandbuch – Auswahl von Lernplattformen“ bearbeitet worden).

Es befinden sich drei Dateien zum download auf dem Server:

- 1) Skriptum "Beurteilungskriterien für Lernsoftware"
(Beurteilung_Lernsoftware.pdf)
- 2) checkliste de.pdf (die Checkliste auf einem DIN-A4-Blatt)
- 3) checkliste de.xls (die Checkliste als Excel-Tabelle)

Achtung: Die Problematik von Checklisten generell ist nicht unerheblich. Ich verweise diesbezüglich auch auf die Arbeiten von PETER BAUMGARTNER.

Bei der ständig steigenden Anzahl von Multimedia-Systemen nimmt auch der Anteil multimedialer Lernsysteme immer mehr zu. Erwartungsgemäß versprechen die Hersteller überschwänglich gesteigerte Lernerfolge und erhöhte Motivation der Lernenden. Um sich nicht auf zweifelhafte Statistiken und subjektive Einschätzungen anderer verlassen zu müssen, ist es wichtig zu wissen, welchen **Kriterien** Software entsprechen muss: *Was ist bei der kritischen Beurteilung von Lernsoftware zu beachten?* Das ist aber nicht nur eine brennende Frage, die Benutzer betreffend. Auch Förderer, Sponsoren und Auftraggeber (Kunden) möchten wissen, nach welchen Kriterien Lernsoftware möglichst objektiv und möglichst *rasch* zu beurteilen ist.

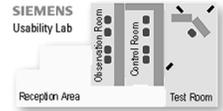


Bild 1.1 Schema des Siemens-Usability-Labs in Johnson City (USA), [W1]

Die Evaluierung von Multimedia-Systemen dient dazu, die Usability eines User-Interfaces zu testen – und zu verbessern.

Nach PREIM (1999) ist die Evaluierung von User-Interfaces wesentlich aufwendiger als z.B. das Testen der technischen Funktionalität von Software.

Wenn wir im Zusammenhang von Multimedia-Systemen über Evaluierung sprechen, müssen wir stets bedenken, dass das System selbst aus einer **Plattform** besteht und einen **Inhalt** (Content) besitzt. Bei der Evaluation muss deshalb einerseits *sowohl* das System *als auch* der Content betrachtet werden, andererseits muss aber bei einigen Kriterien auch differenziert werden: *Was ist System, was ist Content?*

Multimedia-Systeme sind natürlich nur eine Untermenge von Software-Systemen, deswegen gelten naturgemäß auch die Regeln (Ziele, Kriterien und Mittel) der Software-Evaluation. Während allerdings die Ziele und Mittel der Evaluation *anwendungsunspezifisch* sind, müssen die **Kriterien** speziell für Lernsysteme betrachtet werden. Zur ernsthaften Untersuchung von Software solcher Art muss ein Basiswissen über Wissensvermittlung und damit über Kommunikation bekannt sein (siehe Basiswissen Multimedia, Band 2).

Leserinnen und Leser eines solchen Skriptums erwarten etwas ganz Bestimmtes: eine möglichst universelle „Formel“, die *rasch* und *effizient* Ergebnisse bringt – die aber *bei Bedarf* auch tiefer gehende Informationen anbietet. Aus diesem Grund ist dieses Skriptum in *drei Kapitel* gegliedert:

- Kapitel 1: Grundlagen (Definitionen, Arten und Mittel der Software Evaluation),
- Kapitel 2: Checklisten (zuerst ein Überblick über Checklisten und dann *eine konkrete, verwendungsfähige* Checkliste – auch als Excel-Tabelle am Web) und
- Kapitel 3: Erläuterung der Checkliste mit Hinweisen zu den einzelnen Punkten.



Bild 1.2 Ein Blick in den Testraum des Siemens-Usability-Labs [W1]

1 Grundlagen

1.1 Definition Evaluation

Mit Evaluation (sachgerechte Beurteilung) werden zusammenhängende Prozesse der **Wertung** bezeichnet, wie die

- **Auswertung** erhobener Daten und die
- **Bewertung** dieser Daten.

Evaluation ist nicht bloß die Prüfung der Funktion oder die Fehlererkennung. Ziel ist eine Aus- und Bewertung von Einsatzmöglichkeiten und eine **Analyse** der Effekte, die mit dem System bei der **Interaktion** mit Menschen erreicht werden.

Evaluation geht damit weit über eine bloße Messung bzw. quantifizierte (zahlenmäßige) Erfassung einzelner – isolierter – Variablen hinaus. FRICKE (1997) nennt drei wesentliche Bestimmungselemente für „Evaluation“:

systematisch

- Grundlage der Evaluation ist eine **systematisch** gewonnene Datenbasis über Voraussetzungen, Kontext, Prozesse und Wirkungen einer *praxisnahen* Maßnahme.

methodisch

- Evaluation enthält eine Bewertung. Die *methodisch* gewonnenen Daten werden auf dem Hintergrund von **Wertmaßstäben** unter Anwendung bestimmter Regeln *gewichtet*.

prozessorientiert

- Evaluation bezieht sich nicht primär auf die Bewertung des Verhaltens (z.B. Leistungen) einzelner Personen, sondern sollte fester *Bestandteil* der Entwicklung, Realisierung und Kontrolle sein.

Evaluation ist die (analytisch begründete) Einschätzung der Wirkungsweise und Wirksamkeit eines Multimedia-Systems.

Nach Dix u.a. (1998) hat die Software-Evaluation drei Hauptziele:

- Ermittlung von Problemen der Benutzer im Umgang mit der Software,
- Bewertung der Funktionalität der Software und
- Ermittlung der Effekte einer Software auf die Benutzer.

Es besteht ein enger Zusammenhang zwischen dem so genannten Prototyping und der Evaluation in der Softwareentwicklung. Prototyping heisst, dass die Software nicht – wie beim Wasserfallmodell – in einem einzigen, sequentiellen Entwicklungsprozess erstellt wird, sondern dass schon sehr früh arbeitsfähige Programmversionen (Prototypen), gebaut und evaluiert werden.

1.2 Arten von Evaluation

Bei der Beurteilung von Software wird zunächst generell zwischen zwei Arten unterschieden:

- **formative** Evaluation (Überprüfung *während* der Software-Entwicklung) und
- **summative** Evaluation (Bewertung eines bereits bestehenden Software-Produkts)

Bezogen auf den **Software-lifecycle** (siehe Basiswissen Multimedia, Band 3, Modul 1) können die formative und summative Evaluation folgendermaßen dargestellt werden:

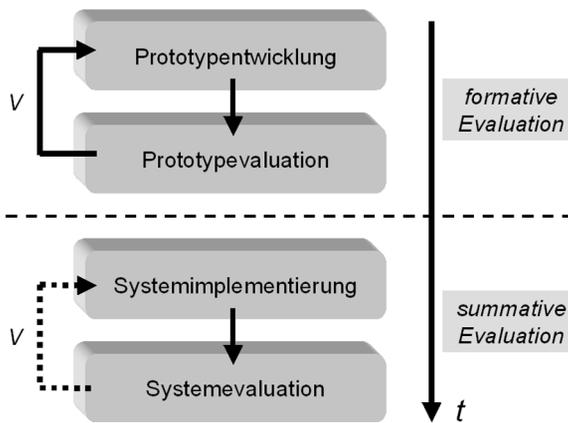


Bild 1.3 Formative und summative Evaluation während des „Software-lifecycle“

Abhängig vom Ziel, das mit einer Evaluation verfolgt wird, ergeben sich unterschiedliche **Kriterien**, die in die Evaluation einfließen. Diese legen fest, *was* und in welchem *Umfang* evaluiert werden soll.

Beispiele für Kriterien sind z.B. Attribute wie Akzeptanz, Erlernbarkeit, Fehlertoleranz, Persönlichkeitsförderlichkeit, Handhabbarkeit, Wartbarkeit, Motivation usw. Ein beliebtes Kriterium, um speziell die Benutzerfreundlichkeit von Software zu „messen“, ist die (subjektive!) Zufriedenheit der Benutzer. Es ist selten möglich, alle Kriterien gleich gut zu erfüllen. Oft müssen Abstriche gemacht werden. Je nach Art der Anforderungen sind einige Kriterien wichtiger als andere.

Während die Frage des **Evaluationszieles** „*Wozu soll evaluiert werden?*“ lautet, heißt sie bei den **Evaluations-Kriterien** „*Was soll evaluiert werden?*“.

Weiterführende Literatur z.B. OPPERMANN u.a. (1992).

Natürliche Anforderungen an Software sind: Fehlerfreiheit, Konsistenz, Standardisierung, Portabilität und Integration

1.3 Mittel der Software-Evaluation

Mit Hilfe der Evaluationsmittel werden die einzelnen Evaluationskriterien zur Erreichung des gewählten Evaluationsziels beurteilt. OPPERMANN (1992) unterscheidet die Evaluationsmittel nach dem Grad dessen, inwieweit bei der Beurteilung der Evaluationskriterien *die Menschen beteiligt sind, die das System auch benutzen*. Darauf aufbauend, klassifiziert er die Evaluationsmittel:

1) Subjektive Evaluationsmittel: Die Evaluationskriterien werden von den Menschen beurteilt, die das System auch benutzen. Unterschieden wird hier zwischen *drei* Möglichkeiten:

- **mündliche Befragung:** Den Benutzern werden Fragen gestellt, die zur Beurteilung der gewählten Evaluationskriterien dienlich sind;
- **schriftliche Befragung:** Ein Fragebogen wird erstellt und von den Benutzern (am Computer oder auf Papier) ausgefüllt; und
- **lautes Denken:** Die Benutzer werden aufgefordert, während einer Aufgabenbewältigung ihre Überlegungen, Probleme und ihre jeweiligen Handlungsalternativen *laut* vor sich her zu sagen.

2) Objektive Evaluationsmittel: Hier wird versucht, möglichst alle subjektiven Einflüsse der Benutzer wie z.B. Emotionen, Vorlieben und Vorurteile weitgehend auszuschließen. Unterschieden wird hier zwischen der

- **anwesenden Beobachtung:** Der Beobachter sitzt neben (oder hinter) dem Benutzer und versucht, anhand der beobachteten Handlungen, Fehler, Ausführungszeiten (Messung) und anderer wahrgenommener Attribute die Evaluationskriterien entsprechend zu beurteilen; und der
- **abwesenden Beobachtung:** Hierbei wird der Benutzer *indirekt* beobachtet (z.B. durch eine Videoaufzeichnung oder logfile-recording).

3) Leitfadenorientierte Evaluationsmittel: Hier wird das Produkt möglichst objektiv entlang eines **Prüfleitfadens** beurteilt: Der Prüfleitfaden ergibt sich aus einer *typischen Aufgabe* des Multimedia-Systems. *Während* der Aufgabenbewältigung wird entsprechend der Evaluationskriterien beurteilt. Für leitfadenorientierte Evaluationsmittel ist *kein* Benutzer notwendig.

Der Zusammenhang zwischen Evaluationszielen, Evaluationskriterien und Evaluationsmitteln ist damit hierarchisch geordnet: Für ein bestimmtes Ziel sind bestimmte Kriterien notwendig und dafür wiederum bestimmte Mittel.

**cognitive
walkthrough =
Analyse korrekter
Handlungsabläufe
und Überprüfung, ob
diese im System
wieder zu finden
sind**

1.4 Skalen

Bei der Erfassung (der Messung) können die Messwerte auf verschiedenen Skalen liegen:

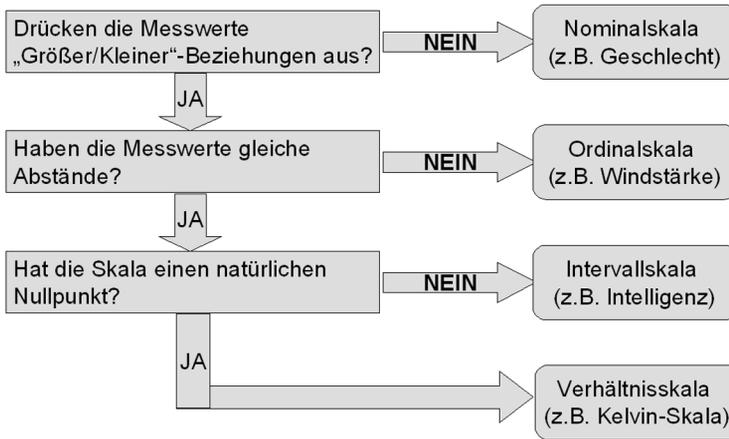


Bild 1.4 Übersicht über die gebräuchlichsten Skalenarten

Unter bestimmten Voraussetzungen können Skalen ineinander übergeführt werden (Skalentransformation).

Beispielsweise können Nominalskalen durch immer differenziertere Klassenbildung in Ordinalskalen (Rangskalen) umgewandelt werden. Umgekehrt können Ordnungen auf Verhältnisskalen zu qualitativen Klassen reduziert werden (z.B. Lebensalter in die Aufteilung: Kinder und Erwachsene).

1.5 Interpretation von Messwerten

Unabhängig von der Frage der verwendeten Skala interessiert die Richtigkeit und die Güte der Messung. Für diese Bewertung sind drei Kriterien wichtig:

- **Reliabilität** (Zuverlässigkeit) als Genauigkeit der Messung: Eine Messung ist **reliabel** (zuverlässig), wenn die Messwerte unabhängig vom Zeitpunkt der Messung sind (verschiedene Messungen mit demselben Messinstrument an denselben Versuchspersonen sollen dasselbe Messergebnis bringen);
- **Validität** (Gültigkeit) als Ausmaß, mit dem eine Messung *tatsächlich* das misst, was es messen soll: Eine Messung ist dann **valide** (gültig), wenn sie das zu messende Merkmal tatsächlich repräsentiert;
- **Objektivität** (Sachlichkeit) als Ausmaß vergleichbarer und kontrollierbarer Umstände: Eine Messung ist dann **objektiv** (sachlich), wenn die erhaltenen Messwerte unabhängig vom Messenden sind.

Messungen sollen sich stets durch möglichst hohe Reliabilität, hohe Validität und hohe Objektivität auszeichnen.

2 Checklisten

2.1 Kriterienkataloge

Checklisten gibt es seit dem Beginn des Einsatzes von Lernsoftware (Educational Software). Allerdings lag und liegt der Fokus stets auf technologischer Seite.

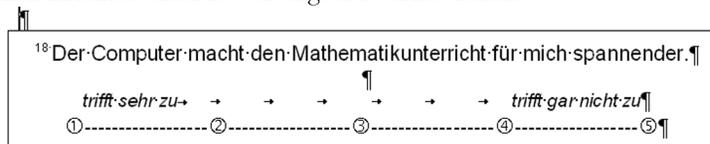
Für summative Evaluationen vielfach genutzte Hilfsmittel stellen so genannte **Kriterienkataloge** dar: Anhand einer *geordneten Liste verschiedener Qualitätskriterien* kann Lernsoftware beurteilt und (numerisch quantitativ) bewertet und verglichen werden.

Allerdings gibt es eine nahezu unüberschaubare Menge an verschiedensten Kriterienkatalogen, und diese sind meistens so detailliert und schwierig anwendbar, dass sie nur von *geschulten Experten* benutzt werden können. SCHREIBER (1998) stellt drei vor:

- **CCEM** (Computer-Based Courseware Evaluation Model) von PRITCHARD, MICCERI & BARRET (1989): 300 Kriterien, gegliedert in 9 Hauptbereiche. Nur ein mehrtägig geschulter Experte kann CCEM anwenden;
- **MEDA** (Méthodologie d'Evaluation des Didacticiels pour Adultes), von der EU ausgearbeitet: ein hierarchisch gegliedertes System von Ja/Nein-Fragen, die abhängig von der Evaluationsart (Entwicklung oder Anwendung) wahlweise abgearbeitet werden können;
- **CES** (Course Evaluation System) von ELLIS, KNIRK, TAYLOR & McDONALD (1993): basiert auf einem empirisch validierten Instruktionsmodell.

Die einzelnen Punkte eines Kriterienkatalogs werden meist in Frageform dargestellt. Es werden wie in jeder Software-Evaluation „qualitative“ und „quantitative“ Daten erhoben. Fragen nach Hersteller, Systemkonfiguration oder Betriebssystem dienen beispielsweise der Abfrage qualitativer Daten. Quantitative Daten werden dagegen meist mittels Rating-Skalen ermittelt. Die Aufgabe des Evaluators besteht vor allem darin, die quantitativen Kriterien innerhalb einer Bewertungsskala einzuordnen.

Bild 1.5 Beispiel für den Entwurf einer Bewertungsskala (Ordinalskala) mit Word



Nach BAUMGARTNER (2002) haben Kriterienkataloge folgende Vorteile:

- Kostengünstig (eine Person evaluiert ein Programm);
- Einfach in der Organisation (Bewertung kann separat vom Einsatzort durchgeführt werden) und diese
- erscheinen vordergründig objektiv.

Allerdings warnt BAUMGARTNER (2002) vor den folgenden Nachteilen:

- Unvollständigkeit;
- Fehlende bzw. strittige Bewertungs- und Gewichtungsverfahren (und damit fehlende Objektivität);
- Theoretische Orientierungslosigkeit (z.B. die Vernachlässigung der zugrunde liegenden didaktischen Konzepte).

2.2 Beurteilungsverfahren

An eigentlichen Beurteilungsverfahren (Zuweisung von Werten) lassen sich grundsätzlich vier bzw. fünf Methoden unterscheiden, die untereinander frei kombinierbar sind:

- **Grading** (Einstufung): Die Evaluanden werden anhand eines vordefinierten Beurteilungsmaßstabes eingestuft;
- **Ranking** (Reihung): Für eine Reihung werden die Evaluanden zueinander in Relation gesetzt und beurteilt. Es entsteht eine Ordinalskala mit einer Aussage wie gut – besser – am besten, die jedoch keine Aussage zu den Abständen untereinander macht;
- **Scoring** (Punktevergabe): Im Unterschied zum Ranking werden hier Punkte vergeben, deren Abstände untereinander gleich und „bedeutungsvoll“ sind. Nur mit dieser Methode sind *summative Operationen* wie Addition oder Division zulässig;
- **Apportioning** (Aufteilung, Zuteilung): Entsprechend der Wertigkeit eines Evaluanden werden vorhandene Ressourcen aufgeteilt (als Beispiel kann die Aufteilung eines Staatshaushaltes genommen werden) und
- **Timing** (Zeitablauf).

STUFFLEBEAM (1972) differenziert in seinem CIPP-Modell (Context, Input, Process, and Product) folgende Evaluationsformen:

- **Kontextevaluation** (Context): Evaluation der Rahmenbedingungen, (z.B. vorhandene Hilfsmittel, Kursinhalt auf Teilnehmer abgestimmt),
- **Inputevaluation** (Input): Evaluation der Teilnehmervoraussetzungen,
- **Prozessevaluation** (Process): Evaluation des Lernprozesses (diese wird auch als **formative Evaluation** bezeichnet) und
- **Produktevaluation** (Product): Evaluation der Lernergebnisse (wird auch als **summative Evaluation** bezeichnet).

Eine quantitative Beurteilung kann in den einzelnen Teilbereichen durch Schulnoten erfolgen: 1 (sehr gut bzw. sehr zufriedenstellend), 2 (gut), 3 (befriedigend), 4 (genügend), 5 (nicht genügend bzw. nicht zufriedenstellend). Oft wird es auch umgekehrt gemacht: Höher ist besser.

Jede quantitative Bewertung soll aber zusätzlich auch qualitativ kommentiert werden, hier vorerst nur durch möglichst objektive Beschreibung – die persönlichen, subjektiven Meinungen können voll in einer „Kür“, einem (subjektiven) Schlussbericht (in Form eines Gutachtens) eingebracht werden.

Der Vorteil der Erfassung qualitativer Punkte liegt in einer schnellen Vergleichbarkeit unterschiedlicher Systeme.

2.3 Checkliste für Multimedia

Auf der folgenden Seite ist eine Checkliste mit *einigen* wichtigen Beurteilungskriterien zu finden. Damit diese Tabelle in der täglichen Praxis zur schnellen Beurteilung von Multimedia-Systemen eingesetzt werden kann (ohne theoretischen Overhead), ist diese Tabelle so gestaltet, dass diese auf einer Seite Platz hat und leicht kopiert werden kann. Sie steht außerdem auf der Website als Excel-Sheet und als pdf zum Download zur Verfügung.

Die Fragen sind in Gruppen angeordnet:

- A ... Allgemeines;
- T ... Technologie;
- H ... Human-Computer Interaction;
- M ... Multimedia;
- L ... Lernen und Didaktik;
- K ... Kognition.

Im Kapitel 3 erfolgt eine systematische Diskussion der einzelnen Punkte dieser Checkliste.

	<i>Beurteilungskriterium</i>	<i>Beurteilung</i>
	<i>(1 = sehr gut; 5 = nicht zufrieden stellend)</i>	<i>(1 ... 5)</i>
A01	Geeignet für diese Zielgruppe	
A02	Preis-Leistungs-Verhältnis	
T01	Ladezeiten (vor allem bei WBT!)	
H01	Bildschirmaufteilung	
H02	Screendesign	
H03	Farben	
H04	Helligkeit	
H05	Formkonstanz	
H06	Informationsinhalt	
H07	Bedienerfreundlichkeit	
H08	Textgröße	
H09	Textlesbarkeit	
H10	Symbolerkennung	
M01	Grafikqualität	
M02	Klangqualität	
M03	Sprachqualität	
M04	Animationsqualität	
M05	Filmqualität	
M06	Interaktivität	
L01	Motivation	
L02	Arousal (Anregung)	
L03	Lerneffizienz	
L04	Lernatmosphäre	
L05	Lernintensität (Reproduktion versus Problemlösen)	
L06	Neugierde	
L07	Überraschung	
L08	Humor	
L09	Erwartungshaltung	
L10	Kreativität	
L11	Schwierigkeit	
L12	Feedback	
L13	Lernzielklarheit	
L14	Lernzielerreichbarkeit	
L15	Übungsmöglichkeit	
L16	Abwechslung	
K01	Wissensvermittlung	
K02	Inzidentielles Lernen	
K03	Problemlösekompetenz	
	Gesamturteil:	

3 Erläuterung der Checkliste

Wichtig ist generell stets ein streng **analytisches Vorgehen**. Am Anfang jeder Evaluierung sollte man mit einer **Fragestellung** beginnen, z.B.:

F: Entspricht die Schriftgröße den Design-Regeln?

Daraus wird eine **Hypothese** gebildet:

H: Die Schriftgröße entspricht den Design-Regeln.

Jetzt wird – aufgrund verschiedener Kriterien – versucht, diese Hypothese zu **verifizieren (zu beweisen)** oder zu **falsifizieren (zu widerlegen)**, z.B.:

Ja, Hypothese H wurde verifiziert: Aufgrund der Untersuchungen und eines Vergleichs mit den geltenden Normen entspricht die Schriftgröße den Design-Regeln.

Oft kann eine Hypothese weder verifiziert noch falsifiziert werden: dann kann **keine Aussage** getroffen werden.

3.1 Allgemeine Beschreibung

Die Beschreibung von Seiten der Hersteller ist oft überzogen – was durchaus verständlich ist, weil ja Kaufentscheidungen bei der Zielgruppe ausgelöst werden sollen. Deshalb ist gerade diese Beschreibung äußerst kritisch zu hinterfragen. Ein erster Schritt einer Evaluation kann sein, dass das Multimedia-Programm in eine so genannte **Hauptströmung** (Baumgartner (1994)) eingeordnet wird. Eine weitere Klassifikation kann nach TAYLOR (1980) durchgeführt werden (vgl. dazu auch Basiswissen Multimedia, Band 2, Modul 5, Kapitel 2):

- **Computer als tutor:** Damit das Programm als Tutor in einem bestimmten Fachgebiet funktioniert, muss dieses von einem Experten programmiert werden. Während der Benutzung treten die Benutzer in einfachen Dialog mit dem Programm, das das Lernmaterial präsentiert, Fragen stellt und evaluiert.
- **Computer als tool:** Damit das Programm als „Werkzeug“ verwendet werden kann, muss dieses eine spezielle Benutzeroberfläche bereitstellen (z.B. Textverarbeitung, Statistikpaket, Tabellen usw.) in der die Studierenden ihre Aufgaben erledigen können (vgl. z.B. Mathematica für den Mathematikunterricht).
- **Computer als tutee:** Um den Computer als „tutee“ zu verwenden – was heißt, dass die Studierenden dem Computerprogramm etwas „beibrin-

Behaviorismus
Kognitivismus
Konstruktivismus

gen“ sollen – , ist es erforderlich, dass die Studierenden die „Programm-
sprache“ beherrschen (vgl. z.B. PROLOG oder LOGO) oder zumindest
mit den Eingabeoperationen sehr vertraut sind (z.B. Authorware oder
Hypercard).

Es sollte stets rasch und sicher erkannt werden, nach welchem
Lernparadigma die Software ausgerichtet ist.

Eine schnelle Einteilung kann folgendermaßen getroffen werden (vgl. mit
Baumgartner (1994):

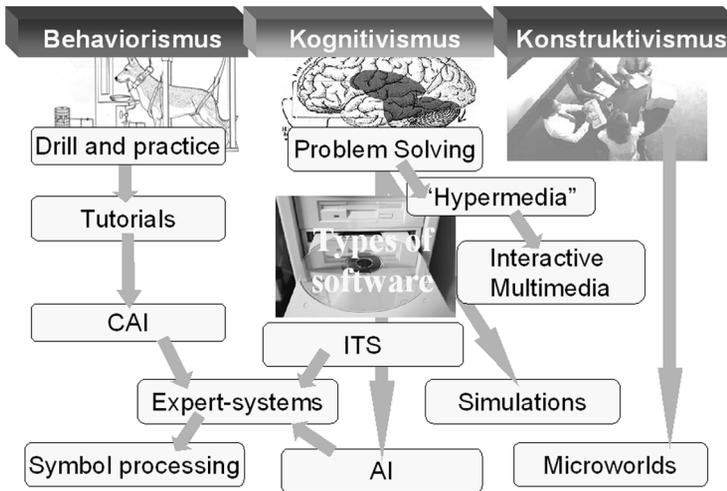


Bild 1.6 Einteilung in Hauptströmung: Welcher Hauptströmung ist das Multimedia-System zuzuordnen und welcher Einzelposition innerhalb dieser Hauptströmung? Geht es in ein Extrem?

Nun zur Erläuterung der einzelnen Beurteilungskriterien in der Checkliste.

3.1.1 A01 Zielgruppe

Oft weicht die empfohlene Zielgruppe (siehe Basiswissen Multimedia, Band 2, Modul 5, Kapitel 3.3.1) von der wirklich *geeigneten* Zielgruppe erheblich ab.

F: Ist die Software für diese Zielgruppe geeignet?

H: Die Software ist für diese Zielgruppe geeignet.

Charakteristika der Zielgruppe können Faktoren wie soziodemographische Daten, Vorwissen, Motivation, Einstellung und Erfahrungen, Lerngewohnheiten, Lerndauer, Lernort usw. sein.

Ein wichtiger Punkt hierbei stellt der Begriff „computer literacy“ dar, was oft schnell übersetzt mit „Vertrautheit mit dem Computer“ umschrieben wird.

3.1.2 A02 Preis-Leistungs-Verhältnis

Das beste Softwareprodukt nutzt den Benutzern nichts, wenn es entweder (für die Zielgruppe) zu teuer oder gar nicht (mehr) erhältlich ist, was leider sehr häufig der Fall ist.

F: Stimmt das Preis-Leistungs-Verhältnis dieser Software?

H: Das Preis-Leistungs-Verhältnis dieser Software stimmt.

Auch (oder gerade bei) Lernsoftware ist eine hohe Leistung zu einem akzeptablen Preis anzustreben. Eine grobe Abschätzung zum Aufwand, der hinter Lernsoftware steht, geben BRUNS & GAJEWSKI (1999):

$$A(MT) = 27,1 MT + 0,4 MT \cdot ST + 1,1 MT \cdot SG + 61,7 MT \cdot hA + 103,7 MT \cdot hV$$

wobei:

A(MT) ... Erstellungsaufwand in Arbeitertagen (früher Mann-Tage)

ST ... Anzahl der Seiten Text

SG ... Anzahl der Seiten Grafik

hA ... Anzahl der Stunden Animation

hV ... Anzahl der Stunden Video

Als Grundlage dient folgender Zeitbedarf:

0,4 MT für eine Seite Text

1,1 MT für eine Seite Grafik

61,7 MT für eine Stunde Animation

103,7 MT für eine Stunde Video

Grundlage dieser Schätzgleichung ist die Auswertung von Daten aus 72 zueinander ähnlichen Lernprogrammen nach einer Studie von WITTE (1995).

3.1.3 Technologischer Hintergrund

Es ist wichtig zu wissen, ob das Programm z.B. unter den Betriebssystemen Windows 3.x, 95, 98, ME, 2000, NT4.0, XP, OS/2, Unix, Linux, Apple usw. läuft.

3.1.4 T01 Systemressourcen

So trivial es klingen mag, aber nicht unerheblich sind die notwendigen Systemressourcen, wie z.B. Prozessor (Pentium I, II, III, IV usw.), verfügbarer Hauptspeicher (z.B. 256 MB RAM) oder Festplattengröße. Ein Beispiel dafür ist das Volksschulprogramm „Lilos Lesewelt“, das enorme Systemressourcen benötigt – aber welche Volksschulklasse verfügt schon über modernste Computer.

F: Decken sich die Ressourcen des verfügbaren Systems mit den Anforderungen?

H: Die Ressourcen des verfügbaren Systems decken sich mit den Anforderungen.

3.1.5 T02 Ladezeiten

Ladezeiten sind heute bei CD-basierenden Produkten kein Problem mehr. Im Gegenteil, ältere Programme können bei zu schnellen CD-Laufwerken Probleme verursachen. Bei web-basierender Software sind die Ladezeiten abhängig von der jeweiligen Netzverbindung (siehe Basiswissen Multimedia, Band 1, Modul 6, Kapitel 2) und können sich erheblich auswirken.

Beispielsweise wird für ein LDTV (Low Definition TV) mit niedriger Bildauflösung (im MPEG-2-Standard) für 1 Stunde Film (675 MB) bei einer ISDN-Verbindung (64 kbit/s) eine Übertragungszeit von 23 Stunden (!) benötigt.

F: Ist die Ladezeit zufrieden stellend?

H: Die Ladezeit ist zufrieden stellend.

3.2 HCI-Aspekte

Vor gar nicht allzu langer Zeit nutzten nur Spezialisten Computer. Das hat sich in den letzten Jahren geändert, sodass Aspekte der Mensch-Maschine Kommunikation (engl. **Human-Computer Interaction**) immer wichtiger wurden. NORMAN (1992) beschreibt in seinem Buch „The Psychology of everyday Things“ das oft unklare Design alltäglicher Dinge. Er definierte zwei Grundprinzipien für gute HCI:

- **Visibility** (Sichtbarkeit; alle „Bedienelemente“ sollen sofort ersichtlich sein);
- **Affordance** (Sinn; bereits das Design sollte den Nutzen bzw. den Mehrwert für die Benutzer erkenntlich machen).

Benutzer wollen keine Features, sondern klare Mehrwerte!

3.2.1 H01 Bildschirmaufteilung

Jede **Bildschirmmaske** (siehe Basiswissen Multimedia, Band 3, Modul 2, Kapitel 2.6) muss so aufgebaut sein, dass sich gleichartige Informationen möglichst an der selben Stelle des Bildschirms befinden. Nach DIN 66 290/1 müssen Masken in vier Informationsklassen aufgeteilt sein: Informations-, Verarbeitungs-, Steuerungs- und Meldungsteil im Größenverhältnis von etwa 1 : 9 : 1 : 2. Der strenge Formalismus nach DIN kann „gelockert“ werden, indem grob in *Arbeitsbereich*, *Orientierungsbereich* und *Steuerungsbereich* eingeteilt wird.

Im **Arbeitsbereich** laufen die lokalen Interaktionen (meistens in Frames) ab. Dort erfolgt die *Vermittlung* der Information (content).

Im **Orientierungsbereich** sollte stets eine *Übersicht* über den aktuellen Standort innerhalb des Lernprogrammes gegeben werden (Wo befindet sich der Benutzer gerade innerhalb des Programms?).

Der **Steuerungsbereich** schließlich dient explizit zur *Navigation* durch das Programm, allenfalls für *Hilfswerkzeuge* (Wörterbuch, Druckfunktion usw.). Nach der *Aufmerksamkeitsverteilung* (im westlichen Kulturkreis) sollte das Erscheinungsbild nach Bild 2.14 (siehe Basiswissen Multimedia, Band 3, Modul 2), gestaltet werden.

F: *Erfolgt die Bildschirmaufteilung nach ergonomischen Gesichtspunkten?*

H: *Die Bildschirmaufteilung erfolgt nach ergonomischen Gesichtspunkten.*

3.2.2 H02 Screendesign

Screendesign stellt im Prinzip die Gestaltung der gesamten Bildschirmoberfläche dar

Mehr als die bloße Aufteilung des Bildschirms in Informationsbereiche ist die Qualität des Screendesigns.

F: *Ist das Screendesign ansprechend?*

H: *Das Screendesign ist ansprechend.*

Designer dürfen niemals vergessen, dass die Bildschirmoberfläche für die Benutzer das *eigentliche System* darstellt (was dahinter liegt – auch wenn es noch so gut ist – ist ihnen egal).

Für ein gutes Screendesign gehört natürlich nicht nur ein ansprechendes, angenehmes und konsistentes (einheitliches) „look & feel“.

Vom kognitionswissenschaftlichen Standpunkt ist es wichtig, dass die Benutzeroberfläche und die Inhalte **gehirngerecht** gestaltet sind, denn die beiden **Gehirnhemisphären** des menschlichen Großhirns sind für unterschiedliche Aufgaben verantwortlich (siehe Basiswissen Multimedia, Band 2, Modul 1, Kapitel 2.2).

Obwohl das Screendesign oft mit einem „schnellen Blick“ durch den Ersteindruck beurteilt werden kann („first impressions are most lasting“), umfasst eine Gesamtbeurteilung des Screendesigns viele Detailspekte – wie nachfolgend besprochen.

3.2.3 H03 Farben

Allgemein unterscheiden wir zwischen zwei Arten von Farben:

- **objektive Farben (Realfarben)** als physikalisch *messbare* und chemisch analysierbare Wirkungen von *Licht* (Basiswissen Multimedia, Band 1, Modul 4, Kapitel 1) und
- **subjektive Farben (Wirkungsfarben)** als *Erregungsqualitäten*, die von verschiedenen Personen *unterschiedlich* empfunden werden.

F: Werden die Farben nach den Gestaltungsprinzipien eingesetzt?

H: Farben werden nach den Gestaltungsprinzipien eingesetzt.

Objektiv optimale Farbkombinationen kann es durch die unendliche Vielfalt von Farben *nicht* geben. Es müssen aber die **Grundprinzipien** der Farbgestaltung eingehalten werden.

Farben lenken die menschliche Aufmerksamkeit sehr stark. Daher sollten Farben grundsätzlich *sparsam* und mit *viel Geschick* eingesetzt werden. Grelle Farben wirken eher hinderlich als förderlich. Außerdem sollten einmal gewählte Farben *konsistent* beibehalten werden.

Im Evaluierungsprozess kann untersucht werden, ob die verwendeten Farben die visuelle Informationsverarbeitung des Menschen wirksam unterstützen, vor allem hinsichtlich:

- Figur-Grund-Unterscheidung (Gestaltpsychologie),
- Gruppierungen von Elementen,
- Suchen, Identifizieren und Zuordnen von Elementen,
- Unterstützung der Erinnerung von Elementen.

Literatur z.B. HELLER (1989), HOLZSCHLAG (2002). EISEMANN (2000).

Gestaltungsprinzipien von Farben, siehe BW MM, Band 3 Modul 3, Kapitel 3

Farbunterschiede werden schneller erkannt als Größen- oder Helligkeitsunterschiede; allerdings werden wiederum Ziffern und Buchstaben schneller erkannt als Farben

3.2.4 H04 Helligkeit

Die Helligkeit (Basiswissen Multimedia, Band 1, Modul 4, Kapitel 1.4.2) jedes Objektes wird immer relativ zu seinem Hintergrund bestimmt. Die wichtigsten Bildschirmfarben haben die Helligkeitsrangfolge Weiß, Gelb, Cyan, Grün, Magenta, Rot, Blau, Braun, Schwarz. Vor einem dunklen Hintergrund sind Weiß, Gelb, Grün und Cyan am besten geeignet. Ältere Benutzer benötigen für Farbumterscheidungen größere Helligkeiten (vgl. mit Zielgruppe).

F: Ist die Helligkeitsverteilung am Bildschirm optimal?

H: Die Helligkeitsverteilung am Bildschirm ist optimal.

3.2.5 H05 Formkonstanz

Aus der Gestaltpsychologie (Basiswissen Multimedia, Band 3, Modul 3, Kapitel 2) ist die Wichtigkeit von Konstanzphänomenen bekannt. Es ergibt sich daraus z.B. die Frage, ob grafische Elemente konsistent (durchgängig) verwendet werden. Plattformkonventionen sollten beachtet werden.

Die Benutzer müssen sich darauf verlassen können, dass gleich aussehende Elemente immer das Gleiche „tun“.

F: Werden grafische Elemente konsistent verwendet?

H: Grafische Elemente werden konsistent verwendet.

3.2.6 H06 Bedienerfreundlichkeit

Wichtig ist vor allem eine *intuitive* Bedienbarkeit. Lange Einarbeitungszeit in die Programmbedienung ist (nicht nur) für Lernprogramme abzulehnen.

Im „worst-case“ würden die Lernenden ein Training (vielleicht ein CBT) brauchen, um ihr CBT zu bedienen.

Die Benutzerführung muss in jedem Fall **selbsterklärend** sein und soll keiner langatmigen „Hilfefunktionen“ oder gar eines Benutzerhandbuches bedürfen.

F: Ist die Bedienerfreundlichkeit gewährleistet?

H: Die Bedienerfreundlichkeit ist gewährleistet.

**Benutzerhandbuch
bitte nicht mit Referenzhandbuch
verwechseln!**

In Software-Magazinen werden immer wieder Tests von Lernprogrammen durchgeführt. Eine Software wurde deswegen als „nicht empfehlenswert“ bezeichnet, weil kein Benutzerhandbuch mitgeliefert wird. Jedoch ist zu prüfen, ob für dessen Bedienung ein Benutzerhandbuch überhaupt notwendig ist – wenn nicht, ist es sogar sehr empfehlenswert.

Achtung: Bedienerfreundlichkeit wird im englischen Sprachraum nicht mit Usability bezeichnet (siehe Basiswissen Multimedia, Band 3, Modul 2). Der Begriff **Usability** setzt sich aus zwei Worten zusammen: use (benutzen) und ability (Fähigkeit), wird im deutschen als „Gebrauchstauglichkeit“ bezeichnet und umfasst weit mehr als nur Bedienerfreundlichkeit: In der ISO Norm 9241 wird Usability als das Ausmaß definiert, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um deren Ziele **effektiv** und **effizient** zu erreichen. Dabei geht es beim Maßstab für die Gebrauchstauglichkeit um den Grad der Effektivität, Effizienz und (auch) Zufriedenstellung.

In einem Usability-Test werden, anhand von quantitativen und qualitativen Testmethoden, auch andere, software-ergonomische Faktoren berücksichtigt (z.B. Aufgabenangemessenheit usw.)

3.2.7 H08 Textgröße

Der Leseabstand (siehe Basiswissen Multimedia, Band 3, Modul 3, Kapitel 1.2.4) zur Bildschirmoberfläche sollte *mindestens 70 cm* betragen. Eine Schrift von 12 Punkt ist auf diese Entfernung gut lesbar. Größer als 18 Punkt sollte die Schrift aber nicht sein (Überschrift ausgenommen).

Ein häufiger Fehler ist es, *vielen verschiedene* Schriftgrößen und Schriftarten miteinander zu kombinieren. Es sollten maximal zwei Schriftarten und zwei Schriftgrößen pro Bildschirmdarstellung verwendet werden.

F: Entspricht die Schriftgröße den Design-Regeln?

H: Die Schriftgröße entspricht den Design-Regeln.

3.2.8 H09 Textlesbarkeit

Im Schwarzdruck auf Papier ist Text wie beispielsweise hier in einer Serifenschrift wie Garamond sehr gut lesbar. Auf dem Bildschirm ist jedoch eine serifenlose Schrift wie Verdana oder Tahoma wesentlich besser lesbar (siehe Basiswissen Multimedia, Band 3, Modul 3, Kapitel 1.1).

F: Ist die Textlesbarkeit zufrieden stellend?

H: Die Textlesbarkeit ist zufrieden stellend.

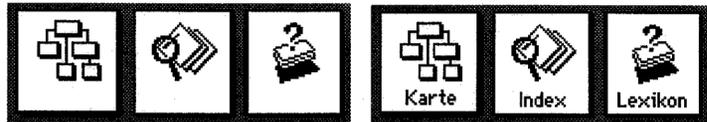
Literatur siehe z.B. ZIEFLE (2002), NEUTZLING (2002), BAINES & HASLAM (2002).

3.2.9 H10 Symbolerkennung

Viele Bildelemente werden durch Icons (Basiswissen Multimedia, Band 3, Modul 3, Kapitel 4.3.7) repräsentiert.

Werden aber zuviele abstrakte Symbole als Icons verwendet, werden die Benutzer vom eigentlichen Lernmaterial stark abgelenkt. Im schlimmsten Fall führt dies zu einem so genannten „cognitive overhead“ (Kognitive Überlast, siehe Basiswissen Multimedia, Band 2, Modul 4, Kapitel 2.4.4) führt. Eine direkte (und simple) Abhilfe kann eine unterlegte Beschriftung von Bedienelementen bringen:

Bild 1.7 Beispiel für Bedienelemente eines Lernprogrammes ohne (links) und mit (rechts) Beschriftung, aus Issing (1997), 283



Nach MOYES & JORDAN (1993) kann man drei Komponenten unterscheiden, mit deren Hilfe Icons bewertet werden können (siehe Basiswissen Multimedia, Band 3, Modul 3, Kapitel 4.4.3, Bild 3.48):

- **Guessability** (Selbsterklärungsgrad) als Maß für den Aufwand der Benutzer, um das erste Mal eine (gewünschte) Funktion auszulösen. Messbar in Zeitdauer, Fehleranzahl und Gesamtzahl der Versuche, bis das Icon richtig interpretiert wird.
- **Learnability** (Erinnerungsgrad) als Maß für den Aufwand der Benutzer, um einen dauerhaften Bedienungskomfort zu erreichen.
- **Experienced User Performance** (EUP) als Maß für den Bedienungskomfort für bereits erfahrene Benutzer; je niedriger der „Überlegungsaufwand“, desto höher der Bedienungskomfort.

F: Erfolgt die Symbolerkennung konsistent nach Grundlagen der Usability?

H: Die Symbolerkennung erfolgt konsistent nach Grundlagen der Usability.

Nach NEGROPONTE (1997) sollten die Benutzer eine „ideale“ Benutzeroberfläche gar nicht erst erlernen müssen. Ja, sie sollten eigentlich gar nicht merken, dass sie in ihren Workflows von einem Computer unterstützt werden.

Literatur z.B. HUANG, SHIEH & CHI (2002), BENBASAT & TODD (1993).

3.3 Multimedialer Hintergrund

In Multimedia-Systemen werden Text, Grafiken (Einzelbilder), Animationen, Audio-material (Sprache, Geräusche, Musik), audiovisuelles Material (Videofilme) eingesetzt.

Die Qualität hängt primär ab von

- der Qualität des **Inhalts** (content),
- der **Didaktik** (*wann* und *wie* wird das Element eingesetzt) und
- der technischen **Performanz** (sind z.B. die Animationen ruckelfrei, ist die Sprache rauscharm usw.).

Zu untersuchen sind dabei – unter anderem – im Detail:

3.3.1 M01 Grafikqualität

Grafik siehe BW MM 3, Modul 3

F: Ist die Grafikqualität ausreichend?

H: Die Grafikqualität ist ausreichend.

3.3.2 M02 Klangqualität

Audio siehe BW MM 3, Modul 4

F: Ist die Klangqualität ausreichend?

H: Die Klangqualität ist ausreichend.

3.3.3 M03 Sprachqualität

Sprache siehe BW MM 3, Modul 4

F: Ist die Sprachqualität ausreichend?

H: Die Sprachqualität ist ausreichend.

3.3.4 M04 Animationsqualität

Animation siehe BW MM 3, Modul 3

F: Ist die Animationsqualität ausreichend?

H: Die Animationsqualität ist ausreichend.

3.3.5 M05 Filmqualität

Video siehe BW MM 3, Modul 3

F: Ist die Filmqualität ausreichend?

H: Die Filmqualität ist ausreichend.

Interaktivität siehe
BW MM 3, Modul 5

3.3.6 M 06 Interaktivität

Gerade die Interaktivität ist ein wichtiges Qualitätsmerkmal der neuen Medien und gerade diese hebt sie deutlich von „*nicht-interaktiven*“, *linearen Medien* (Fernsehen, Film, Video, Dias usw.) ab.

Folgende Unterteilung in der **Art der Interaktion** kann für eine Evaluation hilfreich sein (Die gestellte Grundfrage dabei lautet: Wie erfolgt der Informationszugriff?):

- Vorgegebener Zugriff auf Informationen (passives Rezipieren, wie z.B. Lesen, Zuhören, Anschauen von Animationen usw.);
- freier Zugriff auf Informationen (ebenfalls passives Rezipieren, aber Auswählen, Umblättern, Zurückgehen usw. sind möglich);
- Markierung bestimmter Inhalte und Aktivierung entsprechender Informationen (z.B. durch Anklicken);
- freier Eintrag komplexer Antworten auf komplexe Fragestellungen mit Feedback (vgl. sokratischer Dialog – implementierbar über ein Intelligentes Tutorielles System (ITS), siehe Basiswissen Multimedia, Band 2, Modul 4, Kapitel 3.2);
- freier, ungebundener Dialog mit Tutor oder Lernpartnern (z.B. bei Videoconferencing, chat usw.).

Interaktivität hat große Einflüsse auf die Motivation und sollte daher bei multimedialen Lernprogrammen stets berücksichtigt werden.

F: Ist die Interaktivität zufrieden stellend?

H: Die Interaktivität ist zufrieden stellend.

Fragen, die damit in Zusammenhang stehen, betreffen hauptsächlich die **Kommunikationsform** (passiv – interaktiv – kommunikativ):

- Wie läuft der Dialog zwischen Programm und Benutzer ab?
- Ist es von Seiten der Benutzer nur ein passives Rezipieren?
- Oder gibt es die Möglichkeit, interaktiv oder kommunikativ zu reagieren oder zu agieren (passiv – aktiv)?

Literatur z.B. LUCIN et al. (1998), SAUCIER (2000).

3.4 Didaktik und Lernpsychologie

3.4.1 L01 Motivation

Größere Motivation ist *die* Chance von Multimedia-Systemen.

Gerade bei traditionell unbeliebten Gegenständen, wie beispielsweise Mathematik, kann durch einen didaktisch richtigen Computereinsatz der Gegenstand nicht nur motivierender, sondern auch anschaulicher vermittelt werden. Durch Simulationen können Dinge sichtbar gemacht werden, die von den Lehrenden an der klassischen, traditionellen Tafel einfach nicht dargestellt werden können.

Die brennende Frage ist: *Wie wird Motivation getestet?* Prinzipiell wieder durch die üblichen Erhebungsinstrumente, wie Fragebogen, Interviews, Beobachtung usw. Allerdings sollte dabei eine **Vorerhebung** mit einem **standardisierten Test** durchgeführt werden (z.B. bei 12- bis 18-Jährigen mit dem School Motivation Analysis Test (SMAT) von CATTELL (1970)). So kann ein „Motivationszuwachs“ durch ein Multimedia-System gemessen werden.

Die „Motivierungskraft“ eines Multimedia-Systems hängt primär davon ab, *welche* Motivation bei den Benutzern wirksam wird (siehe Band 2):

- **Erfolgsmotivierte Benutzer** können z.B. durch *ein hohes Maß an Interaktivität* motiviert werden,
- **misserfolgsmotivierte Benutzer** können z.B. durch die *Anonymität am Computer* motiviert werden („*ist ja obnebin nur eine Maschine ...*“), und
- **alle Benutzer** können kurzzeitig durch Auslösung gezielter (lernförderlicher) **Neugier** motiviert werden.

F: Hat dieses Multimedia-System einen hohen Grad an Motivierung?

H: Dieses Multimedia-System hat einen hohen Grad an Motivierung.

3.4.2 L02 Arousal

Das ist ein Punkt, dem allgemein viel zu wenig Beachtung geschenkt wird: Nach BREHM & SELF (1989) ist die **Intensität der Motivation** verbunden mit *Änderungen* im sympathischen Nervensystem (Band 2). Eine höhere Motivation ist auch abhängig von höherem „Arousal“, das in Zusammenhang mit dem „Grad der Wachheit“ steht.

Arousal variiert von sehr niedrigen Werten (Koma oder Schlaf) zu extrem hohen Werten (Panik). Es existiert ein **optimaler „level of arousal“** für die optimalste Lerneffizienz. BERLYNE (1965) zeigte, dass u.a. die wichtigsten

Motivation siehe BW MM, Band 2, Modul 6



Bild 1.8 Die Testzentrale [W4] bietet eine Übersicht über standardisierte Testverfahren

Das Motiv von misserfolgsorientierten Benutzern: Vermeidung von Misserfolg

Arousal, siehe BW MM Band 2, Modul 6

Arousal = Aktivierung

Arousal-Quellen **Stimulation** durch das Lernmaterial, **Sinnhaftigkeit** und die **Neuigkeit** und „**Überraschungen**“ darstellen, die mit oder durch das Lernmaterial oder die Information präsentiert werden.

Die Höhe des **Arousal-Niveaus** ist zudem abhängig von einer „*persönlichen Verantwortung*“, die die Benutzer dem System jeweils widmen (HOLZINGER & MAURER (1999)).

F: Bewirkt dieses Multimedia-System höheres Arousal?

H: Dieses Multimedia-System bewirkt höheres Arousal.

3.4.3 L03 Lerneffizienz

Es handelt sich um den *Unterschied* zwischen angebotenen Material und tatsächlich „behaltener“ Stoffmenge und ist naturgemäß am leichtesten anhand von Faktenwissen (Basiswissen Multimedia, Band 2) messbar. Hier empfehlen sich Messungen an **Vergleichsgruppen**: Eine Gruppe arbeitet mit dem Multimedia-System, die andere traditionell. Dann werden die Ergebnisse verglichen.

F: Führt der Einsatz des Multimedia-Systems zu höherer Lerneffizienz?

H: Der Einsatz des Multimedia-Systems führt zu höherer Lerneffizienz.

3.4.4 L04 Lernatmosphäre

Die Lernatmosphäre hat einen *direkten* Einfluss auf die Intensität der Beschäftigung mit dem Multimedia-System und damit natürlich *indirekt* auch einen Einfluss auf die Lerneffizienz.

Eine Lernatmosphäre ist immer nur subjektiv bewertbar. Manche fühlen sich im „stillen Kämmerlein“ wohl, andere finden die Lernatmosphäre im Kaffeehaus am optimalsten. Entweder man nimmt einen „normierten“ „Normalbenutzer“ oder untersucht die Zielgruppe genau.

F: Ist eine angenehme Lernatmosphäre gewährleistet?

H: Es ist eine angenehme Lernatmosphäre gewährleistet.

Lerneffizienz ist schwierig zu beurteilen

3.4.5 L05 Lernintensität

Ein Hauptkritikpunkt von Multimedia-Gegnern ist, dass durch das „Herumspielen“ die Lernintensität abnimmt (Motto: „TV is easy – books are hard“). Aber durch erhöhte Motivierung, konkretes Anbieten von Faktenwissen *und* Möglichkeiten zum Erwerb von Problemlösewissen kann die Lernintensität durchaus auch erhöht werden.

F: Wird ein hoher Grad an Lernintensität gewährleistet?

H: Es wird ein hoher Grad an Lernintensität gewährleistet.

Die Möglichkeit zur aktiven Lösung von Problemen ist enorm wichtig, denn – wie eine Untersuchung von HOLZINGER (1997) an 722 Versuchspersonen gezeigt hat – steht im traditionellen Schulsystem das Lernen von Fakten und nicht das aktive Problemlösen im Vordergrund (stark vom Schultyp abhängig). Faktenwissen ist zwar als eine Grundlage (Basiswissen) unverzichtbar, in der Praxis wird aber Problemlösewissen benötigt.

3.4.6 L06 Neugierde

Darunter verstehen wir allgemein die aus dem natürlichen Orientierungsverhalten des Menschen ableitbare Tendenz, subjektiv Neues zu erleben und zu erkunden.

Dieser **Explorationsdrang** ist je nach Vorwissen, Alter und Motivation verschieden. Neugierde ist ein wesentlicher Ausgangspunkt für den Grad der Anregung (Arousal).

F: Wird durch das Multimedia-System Neugierde gefördert?

H: Durch das Multimedia-System wird die Neugierde gefördert.

3.4.7 L07 Überraschung

Überraschung fördert Motivation, Arousal und Neugierde.

Deshalb haben „Überraschungseffekte“ einen sehr hohen Stellenwert. Allerdings ist es für Multimedia-Designer natürlich schwierig, „ständige Überraschungseffekte“ einzubauen.

F: Bietet das Multimedia-Lernprogramm einen hohen Überraschungseffekt?

H: Das Multimedia-Lernprogramm bietet einen hohen Überraschungseffekt.

3.4.8 L08 Humor

Ein Aspekt, der ebenfalls sehr oft vernachlässigt wird: Humor wird nach ISEN u.a. (1987) als universeller „Urquell“ für die Produktivität und auch die Problemlösekapazität an Arbeitsplätzen (Teamwork) angesehen.

F: Wird innerhalb des Multimedia-Systems ausreichend Humor geboten?

H: Innerhalb des Multimedia-Systems wird ausreichend Humor geboten.

3.4.9 L09 Erwartungshaltung

Jeder Benutzer hat gewisse Vorstellungen vom und Erwartungen an das Multimedia-System. Es ist natürlich extrem wichtig, dass diese (subjektiven) Erwartungen erfüllt werden. Ansonsten werden die Benutzer enttäuscht.

F: Werden die Erwartungen der Benutzer (Zielgruppe) erfüllt?

H: Die Erwartungen der Benutzer (Zielgruppe) werden erfüllt.

3.4.10 L10 Kreativität

Um Kreativität zu fördern, ist es z.B. erforderlich, dass das System ein hohes Maß an Interaktion zulässt. Es sollte auch genügend **Freiraum** für die Benutzer für *eigene* Versuche, Experimente und der Umsetzung eigener Ideen vorhanden sein.

F: Fördert das Multimedia-System die Kreativität?

H: Das Multimedia-System fördert die Kreativität.

3.4.11 L11 Schwierigkeit

Die Benutzung des Systems sollte prinzipiell nicht schwierig sein. Gemeint ist hier die Schwierigkeit des Fachinhaltes (Content). Diese ist stark vom **Vorwissen** der Benutzer abhängig und wieder *nur subjektiv beurteilbar*: Was von manchen als extrem schwierig empfunden wird, kann von anderen als „lächerlich einfach“ empfunden werden. Bei der Beurteilung dieses Punktes ist die Abstimmung mit dem Punkt **Zielgruppe** zwingend erforderlich.

F: Ist die Schwierigkeit des Contents für die Zielgruppe angemessen?

H: Die Schwierigkeit des Contents ist für die Zielgruppe angemessen.

3.4.12 L12 Feedback

Insbesondere bei Prüfungsfragen ist dieser Punkt sehr wichtig, damit die Benutzer einerseits eine direkte Rückmeldung über ihre Antworten bekommen, andererseits die Neugierde gefördert wird („*bin gespannt auf die nächste Reaktion auf meine Eingabe ...*“).

F: *Gewährleistet das Multimedia-System ausreichendes Feedback?*

H: *Das Multimedia-System gewährleistet ausreichendes Feedback.*

3.4.13 L13 Lernzielklarheit

Lernziele sind in jeder Lehr-/Lernsituation extrem wichtig. Es sollte stets angegeben werden, *was* das eigentliche **Ziel** eines Kurses ist.

F: *Werden Lernziele ausreichend angegeben?*

H: *Lernziele werden ausreichend angegeben.*

3.4.14 L14 Lernzielerreichbarkeit

Es ist andererseits ebenfalls wichtig, dass es für die Benutzer (mit oft sehr unterschiedlichem Vorwissen) auch *möglich ist*, die gesetzten Lernziele zu erreichen. Dazu ist eine **Auswahl der Schwierigkeitsstufe** hilfreich.

**Levels in der Art
„leicht –
mittel –
schwierig“**

Können Lernziele nicht erreicht werden, kann sich extreme Frustration bei den Benutzern ausbreiten, die bis zur totalen Demotivation führen kann. Andererseits können aber dieselbe Wirkung auch zu leichte Inhalte haben.

F: *Sind die Lernziele für alle Benutzer der Zielgruppe erreichbar?*

H: *Die Lernziele sind für alle Benutzer der Zielgruppe erreichbar.*

3.4.15 L15 Übungsmöglichkeit

Es sollte stets die Möglichkeit bestehen, unterschiedliche Übungen zum Lernstoff durchzuführen. Darüber hinaus ist es günstig, wenn die Möglichkeit *zusätzlicher* Übungen für weiter interessierte Benutzer besteht.

F: *Bietet das Multimedia-System ausreichend Übungsmöglichkeiten?*

H: *Das Multimedia-System bietet ausreichend Übungsmöglichkeiten.*

3.4.16 L16 Abwechslung

Abwechslung fördert Motivation und Arousal (siehe dort).

F: *Stellt das Multimedia-System genügend Abwechslung bereit?*

H: *Das Multimedia-System stellt genügend Abwechslung bereit.*

3.4.17 Was ist noch zu beachten?

Ablenkende Elemente. Durch z.B. nicht optimale Auswahl von Farben, falsche Textgrößen, mangelnde Strukturierung und vielen anderen Details gibt es eine Fülle von Möglichkeiten für **Ablenkreize** (Distraktoren), die dem eigentlichen Lernprozess nicht förderlich sind.

Storyboard O.K. Genauso wie bei einer Theater- oder Filmproduktion ist für die Entwicklung von multimedialen Lernsystemen ein Konzept – eine Art „Drehbuch“ – notwendig (siehe Basiswissen Multimedia, Band 3, Modul 4, Dramaturgie).

Aber Vorsicht, multimediale Lernsysteme sind nichtlinear, d.h., sie laufen nicht sequentiell ab wie beim klassischen Film. Die Benutzer sollen ja gerade *interaktiv* den Ablauf des Programms *beeinflussen* können.

Aus diesem Grund wird bei Lernsoftware auch nicht von Drehbuch (script), sondern eben von **Storyboard** gesprochen. Das Storyboard ist der Kern jeder Entwicklung von Lernsoftware – technisch gesehen einem Flussdiagramm sehr ähnlich.

Viele Storyboards orientieren sich am **UCIT (Universal Constructive Instructional Theory)**, was einen interessanten Ansatz darstellt, den gesamten Instrukionsprozess zu *strukturieren*. Seit der Antike – und natürlich insbesondere beim Einsatz multimedialer computergestützter Lernmedien – geht es dabei zentral um das „didaktische Dreieck“ (Bild 1.9).

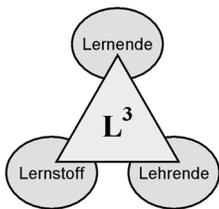


Bild 1.9 Seit der Antike bekannt: das didaktische Dreieck

Die **Dramaturgie** (Verlauf einer Handlung) beeinflusst wesentlich die *Akzeptanz* der Software.

Eine Frage lautet: Wie wirkt das Programm nach mehrmaligem Durchlauf? Das ist ein ganz kritischer Punkt: Oft ist Lernsoftware nur beim erstmaligen Durchlauf spannend und interessant und verliert nach mehrmaligen Durchlauf enorm an Reiz. Das hat natürlich unmittelbare Auswirkungen auf den Lernerfolg.

3.5 Kognitionspsychologischer Hintergrund

3.5.1 K01 Wissensvermittlung

Welche Art von Wissen wird im Multimedia-Programm vermittelt? Wird reines Faktenwissen vermittelt, Konzepte oder vielleicht sogar Strategien zur Problemlösung? Oder wird eine *Kombination* an verschiedenen Wissensarten vermittelt? Faktenwissen stellt die Grundlage für Konzeptwissen dar. Beides wiederum stellt die Grundlage von Problemlösekompetenz dar. Deshalb ist es sehr wichtig, möglichst *vielen Wissensarten* anzubieten (siehe Basiswissen Multimedia, Band 2, Modul 2, Kapitel 3).

Auch Faktenwissen hat – als Basis für Konzeptwissen und Problemlösewissen – große Bedeutung!

F: Wird durch das Multimedia-System Wissen unterschiedlich vermittelt?

H: Durch das Multimedia-System wird Wissen unterschiedlich vermittelt.

Literatur z.B. BALLSTEDT (1997).

3.5.2 K02 Inzidentielles Lernen

Erfolgt das Lernen intentional (also intendiert – mit voller Absicht der Benutzer, etwas zu lernen) oder erfolgt das Lernen im Multimedia-System auch **inzidentuell** (d.h. nicht intendiert – im Idealfall ohne dass die Benutzer explizit beabsichtigen zu „lernen“). Nach HOLZINGER & MAURER (1999) sollte der Anteil an inzidentiellem Lernen nicht unterschätzt werden, letztendlich lernen wir tagtäglich vieles inzidentuell.

Game-Shows haben eine starke inzidentielle Komponente. Es wird kaum jemand leugnen, bei z.B. „*Wer wird Millionär?*“ zu lernen, ohne sich dessen explizit bewusst zu sein.

F: Erfolgt durch das Multimedia-System auch inzidentielles Lernen?

H: Durch das Multimedia-System erfolgt auch inzidentielles Lernen.

3.5.3 K03 Problemlösen

Können in der Lernumgebung aktiv Probleme gelöst werden und wenn ja, in welchem Ausmaß (Problemlösen, siehe Basiswissen Multimedia, Band 2, Modul 2, Kapitel 5). Siehe auch L15 Übungsmöglichkeit.

F: Wird durch das Multimedia-System Problemlösekompetenz gefördert?

H: Durch das Multimedia-System wird Problemlösekompetenz gefördert.

3.5.4 Gesamturteil

Um es auf einen Nenner zu bringen, sollte auch ein Gesamturteil abgegeben werden. Das ist allerdings schwierig und kann selten gerecht sein. Schließlich hat jedes Produkt Vor- *und* Nachteile und kaum eines nur Vorteile (oder auch nur Nachteile). Ein Gesamturteil sollte aber immer hinterfragt werden.

An letzter Stelle sollten die Beurteiler eigene Empfehlungen für mögliche Verbesserungen aus ihrer Sicht abgeben.

Empfehlung: Was könnte wie verbessert werden?

4 Zusammenfassung

Mit Evaluation wird die *Auswertung* erhobener Daten und die *Bewertung* (*Beurteilung*) dieser Daten bezeichnet. Eine Evaluation sollte stets **systematisch**, **methodisch** und **prozessorientiert** durchgeführt werden. Es wird unterschieden zwischen **formativer** Evaluation (während der Entwicklung) und **summativer** Evaluation (nach Fertigstellung). **Subjektive Evaluation** schließt die *mündliche* und die *schriftliche Befragung* und das *laute Denken* ein. **Objektive Evaluation** bedient sich der *anwesenden* und *abwesenden Beobachtung*. Bei leitfadenorientierten Evaluationsmitteln wird das Produkt entlang eines Prüflitfadens beurteilt, der sich aus typischen Aufgaben des Systems ergibt. Bei der *Erfassung* der Messwerte können verschiedene **Skalen** (z.B. Nominalskala, Rangskala, Verhältnisskala usw.) verwendet werden, die unter bestimmten Voraussetzungen durch eine **Skalentransformation** ineinander übergeführt werden können. Messungen sollen sich stets durch hohe **Reliabilität** (Zuverlässigkeit), **Validität** (Gültigkeit) und **Objektivität** (Sachlichkeit) auszeichnen. Als **Beurteilungsverfahren** (*Zuweisung von Werten*) werden **Grading** (Einstufung), **Ranking** (Reihung), **Scoring** (Punktevergabe) und **Apportioning** (Aufteilung, Zuteilung) verwendet. Eine **quantitative Beurteilung** (Vorteil: leichte Vergleichbarkeit von Systemen) kann durch *Schulnoten* erfolgen, aber oft wird es auch umgekehrt gemacht: *Höher ist besser*. Multimedia-Systeme können *systematisch* mit **Checklisten** beurteilt werden. Zu den ersten und *allgemeinsten Fragestellungen* (**A-Fragen**) gehören die Frage nach der *Zielgruppe* und das *Preis-Leistungs-Verhältnis*. Beim technologischen Hintergrund (**T-Fragen**) interessiert außer den *Hardware- und Softwarevoraussetzungen* (für Benutzer wichtig) und der verwendeten *Programmiersprache oder des Autorensystems* (für Designer interessant) vor allem die *Ladezeit*. Zu den HCI-Aspekten (**H-Fragen**) zählen Kriterien wie eine ergonomische *Bildschirmaufteilung* (*Arbeits-, Orientierungs- und Steuerungsbereich*), das *Screendesign*, die Verwendung von *Farben, Helligkeit, Formkonstanz*, *Bedienerfreundlichkeit* (Usability) und Kriterien bezüglich *Textgröße, Textlesbarkeit* und *Symbolerkennung* (Guessability, Learnability und Experienced User Performance). Zum multimedialen Hintergrund (**M-Fragen**) zählen *Grafik-, Klang- und Sprachqualität, Animations- und Filmqualität* und Grad und Art der *Interaktivität*. In den Bereich der Didaktik und Lernpsychologie (**L-Fragen**) fallen Kriterien wie *Motivation, Arousal, Lerneffizienz, Lernatmosphäre, Lernintensität* (Reproduktion versus Problemlösen), *Neugierde, Überraschung, Humor, Erwartungsbaltung, Kreativität, Schwierigkeit, Feedback, Lernzielklarheit, Lernzielerreichbarkeit, Übungsmöglichkeit* und *Abwechslung*. In den kognitionspsychologischen Bereich (**K-Fragen**) gehören *Wissensvermittlung* (Faktenwissen, Konzeptwissen, Problemlösewissen), *inzidentielles Lernen* und *Problemlösekompetenz*. Ein **Gesamturteil** ist sehr schwierig abzugeben und kann selten gerecht sein, weil jedes Produkt ganz unterschiedliche Vor- und Nachteile hat.

**Evaluation =
sachgerechte
Beurteilung,
Bewertung**

5 Literatur

5.1 Bücher

BALLSTAEDT, STEFFEN PETER (1997): Wissensvermittlung: Die Gestaltung von Lernmaterial. Weinheim: Psychologie Verlags Union.

BAINES, PHIL; HASLAM, ANDREW (2002) Lust auf Schrift! Basiswissen Typografie. Mainz: Schmidt.

BAUMGARTNER, PETER; HÄFELE, HARTMUT, MAIER-HÄFELE, KORNELIA (2002): E-Learning Praxishandbuch. Auswahl von Lernplattformen. Innsbruck: Studien-Verlag.

BAUMGARTNER, PETER; PAYER, SABINE (1994): Lernen mit Software. Innsbruck: Studien-Verlag.

BORTZ, JÜRGEN; DÖRING, NICOLA (1995): Forschungsmethoden und Evaluation. 2. Auflage. Berlin, Heidelberg, New York: Springer.

BRICKENKAMP (1997): Handbuch psychologischer und pädagogischer Tests. 2. Auflage. Göttingen: Hogrefe.

DIX, ALAN (1995): Mensch, Maschine, Methodik. München u.a.: Prentice Hall.

EBERLEH, EDMUND; OBERQUELLE, HORST; OPPERMAN, REINHARD, Hrsg. (1994): Einführung in die Software-Ergonomie: Gestaltung graphisch-interaktiver Systeme: Prinzipien, Werkzeuge, Lösungen. Berlin: de Gruyter.

HELLER, EVA (1989): Wie Farben wirken: Farbpsychologie, Farbsymbolik, kreative Farbgestaltung. Reinbek bei Hamburg: Rowohlt.

HERCZEG, MICHAEL (1994): Grundlagen der Mensch-Computer-Kommunikation. Bonn, Paris, Reading (MA): Addison-Wesley.

HOLZSCHLAG, MOLLY E. (2002): Farbe für Websites. Reinbek bei Hamburg: Rowohlt.

HOLZINGER, ANDREAS (2002): Basiswissen Multimedia. Band 1: Technik. Technologische Grundlagen multimedialer Informationssysteme. 2. Auflage. Würzburg: Vogel.

HOLZINGER, ANDREAS (2001): Basiswissen Multimedia. Band 2: Lernen. Kognitive Grundlagen multimedialer Informationssysteme. Würzburg: Vogel.

HOLZINGER, ANDREAS (2001): Basiswissen Multimedia. Band 3: Design. Entwicklungstechnische Grundlagen multimedialer Informationssysteme. Würzburg: Vogel.

- HOLZINGER, ANDREAS (2002): Basiswissen IT/Informatik. Band 1: Informatik-
onstechnik. Würzburg: Vogel.
- HOLZINGER, ANDREAS (2003): Basiswissen IT/Informatik. Band 2: Informatik.
Würzburg: Vogel.
- LIENERT; RAATZ (1994): Testaufbau und Testanalyse. 5. Auflage. Weinheim: PVU.
- NEGROPONTE, NICHOLAS (1997): Total digital. Die Welt zwischen 0 und 1 oder Die
Zukunft der Kommunikation. München: Goldmann.
- NEUTZLING, ULLI (2002): Typo und Layout im Web. Reibek bei Hamburg: Rowohlt.
- OPPERMANN, REINHARD; MURCHNER, BERND, REITERER, HARALD (1992): Soft-
ware, ergonomische Evaluation. Berlin: deGruyter.
- OPPERMANN, R.; MURCHNER, B.; REITERER, H.; KOCK, M. (1992): Softwareergono-
mische Evaluation (Der Leitfaden EVADIS II). Berlin: de Gruyter.
- REINMANN-ROTHMEIER, G.; MANDL, H.; PRENZL, M. (1994): Computerunterstüt-
zte Lernumgebungen: Planung, Gestaltung und Bewertung. Erlangen: Publicis-MCD-
Verlag.
- SCHENKEL, PETER, Hrsg. (1995): Evaluation multimedialer Lernprogramme und
Lernkonzepte. Berichte aus der Berufsbildungspraxis. Nürnberg: Verlag Bildung und
Wissen.
- SCHULMEISTER, ROLF (1997): Grundlagen hypermedialer Lernsysteme: Theorie, Di-
daktilik, Design. München, Wien: Oldenbourg, 387 – 414.
- SCHULMEISTER, ROLF (2003): Lernplattformen für das virtuelle Lernen. Evaluation
und Didaktik. München, Wien: Oldenbourg.
- SCHREIBER, ALFRED (1998): CBT – Anwendungen professionell entwickeln: Pla-
nung, Entwicklung und Bewertung von Lernsoftware für die Praxis. Berlin u.a.:
Springer.
- THOMÉ, D. (1989): Kriterien zur Bewertung von Lernsoftware. Heidelberg: Hüthig.
- TIMPE, K.-P.; JÜRGENSOHN, T.; KOLREP, H., Hrsg (2002): Mensch- Maschine- Sys-
temtechnik. Konzepte, Modellierung, Gestaltung, Evaluation. Symposium Publi-
shing.
- ZIEFLE, MARTINA (2002): Lesen am Bildschirm: eine Analyse visueller Faktoren (In-
ternationale Hochschulschriften; 375). Münster u.a.: Waxmann.

5.2 Artikel

AUHUBER T, SCHULZ S, SCHRADER U, KLAR R. (1997): Ein Modell zur Evaluation medizinischer CBT-Programme. In: Muche R.; Büchele G.; Harder D.; Gaus W., Hrsg. Medizinische Informatik, Biometrie und Epidemiologie – GMDS '97. Medizinische Informatik, Biometrie und Epidemiologie. München: MMV Medizin Verlag 1997; 126–130.

BAUMGARTNER, PETER (2002): Pädagogische Anforderungen für die Bewertung und Auswahl von Lernsoftware. In: Issing, Ludwig J.; Klimsa, Paul, Hrsg. Information und Lehren mit Multimedia und Internet: Lehrbuch für Studium und Praxis. 3. Auflage. Weinheim: Beltz, Psychologie Verlags Union.

BAUMGARTNER, PETER (1997): Evaluation vernetzten Lernens: 4 Thesen. In: Simon, Hartmut, Hrsg: Virtueller Campus. Forschung und Entwicklung für neues Lehren und Lernen, Münster: Waxmann.

FRICKE, R. (1997): Evaluation von Multimedia. in: Issing, L., Klimsa, P., Hrsg.: Information und Lernen mit Multimedia. Weinheim, Basel: Beltz. 401 – 414.

5.3 Books in English

DIX, ALAN J.; FINLAY, JANET E.; ABOARD, GREGORY D.; BEALE RUSSELL (1998): Human-Computer Interaction. Second Edition. London et. al: Prentice Hall.

EISEMAN, LEATRICE (2000): Pantone Guide to Communicating with Color. Cincinnati (OH): Northlight Books.

HANNAFIN, MICHAEL J.; PECK, KYLE L. (1997): The Design Development and Evaluation of Instructional Software. MacMillan.

HOLLANDER, NATHAN (2000): A Guide to Software Package Evaluation & Selection: The R2Isc Method. AMACOM.

LANDAUER, THOMAS K. (1996): The trouble with computers: Usefulness, usability, and productivity. Cambridge (MA): MIT Press.

LINDGAARD, G. (1994): Usability Testing and System Evaluation: A Guide for Designing Useful Computer Systems. London: Chapman and Hall.

MEISTER, DAVID (1986): Human Factors Evaluation and Testing. Amsterdam: Elsevier.

OPPENHEIM, A.N. (1999): Questionnaire Design, Interviewing and Attitude Measurement. London, Washington: Pinter.

PREECE, JENNY; ROGERS, YVONNE; SHARP, HELEN; BENYON, DAVID; HOLLAND, SIMON; CAREY, TOM (1994): Human-Computer Interaction. Wokingham (UK): Addison Wesley.

SAUCIER, CHRISTINE (2000): *Web Animation & Interactivity*. New York: Delmar.

SCRIVEN, MICHAEL (1991). *Evaluation Thesaurus*. 4th Editionl. Newbury Park: Sage Publications.

STUFFLEBEAM, D. L.; SHINKFIELD, A. J. (1990). *Systematic Evaluation*. Boston (MA): Kluwer-Nijhoff.

5.4 Articles in English

ANDERSON, E. E.; CHEN, Y-M. (1997): Microcomputer software evaluation: An econometric model. *Decision Support Systems*, 19 (2), 75 – 92.

BENBASAT, I.; TODD, P. (1993): An experimental investigation of interface design alternatives: icon versus text and direct manipulation versus menus. *International Journal of Man-Machine Studies*, vol. 38, 369-402.

ELLIS, J.A.; KNIRK, F.G.; TAYLOR, B.E.; MCDONALD, B. (1993): The course evaluation system. *Instructional Science*, 21, 313-334.

HOLZINGER, A.; MAURER H. (1999): Incidental learning, motivation and the Tamagotchi Effect: VR-Friends, chances for new ways of learning with computers. *CAL99 Abstract Book*. London: Elsevier, 70.

HUANG, S.-M.; SHIEH, K.-K; CHI, C.-F. (2002): Factors affecting the design of computer icons. *International Journal of Industrial Ergonomics*, vol. 29, 211-218.

JONES, A.; SCANLON, E.; TOSUNOGLU, C.; MORRIS, E.; ROSS, S.; BUTCHER, P; GREENBERG, J. (1999): Contexts for evaluating educational software. *Interacting with Computers*, Volume 11, Issue 5, May 1999, 499 – 516.

LAITENBERGER, OLIVER (2001): Cost-effective detection of software defects through perspective-based inspections, *Empirical Software Engineering*, Volume 6, Issue 1, March 2001, Pages 81-84.

LUCKIN, R.; PLOWMAN, L.; GJEDDE, L.; LAURILLARD, D; STRATFOLD, M.; TAYLOR, J. (1998). An evaluator's toolkit for tracking interactivity and learning. In Oliver, M. ed. *Innovation in the Evaluation of Learning Technologies*, pp.42-64. London: University of North London.

MEMON, ATIF M.; POLLACK, MARTHA E.; SOFFA, MARY LOU (2001): Hierarchical GUI test case generation using automated planning, *IEEE Transactions on Software Engineering*, Volume 27, Issue 2, February 2001, Pages 144-155.

MOYES, J.; JORDAN, P. W. (1993): Icon Design and Its Effect on Guessability, Learnability, and Experienced User Performance. In: Alty, J. L.; Diaper, D.; Guest, S. ,eds., *People and Computers VIII: Proceedings of the Human-Computer Interaction (HCI'93) Conference*, Cambridge University Press, Cambridge, UK, 1993, 49-59.

O'CONNELL, EMILIE; SAIEDIAN, HOSSEIN (2000): Can you trust software capability evaluations? *Computer*, Volume 33, Issue 2, 2000, 28–35.

OLIVER, M. (2000). An introduction to the Evaluation of Learning Technology. *Educational Technology & Society*, 3(4).

SCRIVEN, MICHAEL (1999): The Nature of Evaluation Part I: Relation to psychology. *Practical Assessment, Research & Evaluation*, 6 (11).

SCRIVEN, MICHAEL (2001): Evaluation: future tense. *The American Journal of Evaluation*, Volume 22, Issue 3, Autumn 2001, 301-307.

STAMELOS, I.; VLAHAVAS, I.; REFANIDIS, I.; TSOUKIÀS, A. (2000): Knowledge based evaluation of software systems: a case study. *Information and Software Technology*, Volume 42, Issue 5, 1 April 2000, 333 – 345.

STUFFLEBEAM, D. L. (1972): The relevance of the CIPP evaluation model for educational accountability. *SRIS Quarterly*, 5(1).

SQUIRES, D.; MCDUGALL, A. (1996) Software evaluation: a situated approach. *Journal of Computer Assisted Learning*, 12, 146 – 161.

SQUIRES, DAVID; PREECE, JENNY (1999): Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, Volume 11, Issue 5, May 1999, 467 – 483.

SQUIRES, DAVID; PREECE, JENNY (1996): Usability and learning: Evaluating the potential of educational software. *Computers and Education*, 27(1), 15-22.

TSAGIAS, MANOLIS; KITCHENHAM, BARBARA (2000): An evaluation of the business object approach to software development. *Journal of Systems and Software*, Volume 52, Issues 2 – 3, 1 June 2000, 149 – 156.

ZAHNER, J. E., REISER, R. A., DICK, W. & GILL, B. (1992). Evaluating instructional software: A simplified model. *Educational Technology Research and Development*, 40 (3), 55 – 62.

VOLPE, R. M.; AQUINO, M. T. B.; NORATO D. Y. J. (1998): Multimedia system based on programmed instruction in medical genetics: construction and evaluation. *International Journal of Medical Informatics*, Volume 50, Issues 1-3, June 1998, 257 – 259.

5.5 Journals

Educational Research and Evaluation (ISSN: 1380-3611, electronic:1380-3611) | Swets & Zeitlinger Publishers

Evaluation (ISSN: 1356-3890, electronic: 1356-3890) | SAGE Publications

Instructional Science (ISSN: 0020-4277) | Kluwer

Practical Assessment, Research and Evaluation (ISSN 1531-7714) | ERIC

Quality Assurance in Education (ISSN: 0968-4883) | Emerald Publications

Studies in Educational Evaluation (ISSN: 0191-491X) | Elsevier

The American Journal of Evaluation, formerly Evaluation Practice (ISSN: 1098-2140) | Elsevier

5.6 Internet-Links

Internet-Links zu diesem Skriptum sind großteils auf der Buch-Homepage www.basiswissen-multimedia.at unter Design Modul 6: Evaluation verfügbar!

Danke für Ihre Aufmerksamkeit!