

**Grundlagen der Evaluation
computer- und netzgestützter
Lernangebote**

Research Report 9

Kim, H.O.; Klauser, F.

Herausgeber.:
Bogaschewsky, R.; Hoppe, U.; Klauser, F.; Schoop, E.; Weinhardt, Ch.

Kim, H.O.; Klauser, F

**Grundlagen der Evaluation computer- und netzbasierter
Lernangebote**

IMPULS^{EC} Research Report 9

Herausgeber: Bogaschewsky, R.; Hoppe, U.; Klauser, F.;
Schoop, E.; Weinhardt, Ch.

Die Herausgeber

Prof. Dr. Ronald Bogaschewsky
Julius-Maximilians-Universität Würzburg
Lehrstuhl für BWL und Industriebetriebslehre
D-97070 Würzburg

Prof. Dr. Uwe Hoppe
Universität Osnabrück
Lehrstuhl für BWL/ Organisation und Wirtschaftsinformatik
D-49069 Osnabrück

Prof. Dr. Fritz Klauser
Universität Leipzig
Lehrstuhl für Berufs- und Wirtschaftspädagogik
D-04229 Leipzig

Prof. Dr. Eric Schoop
Technische Universität Dresden
Lehrstuhl für Wirtschaftsinformatik, insb. Informationsmanagement
D-01062 Dresden

Prof. Dr. Christof Weinhardt
Universität Karlsruhe (TH)
Lehrstuhl für Informationsbetriebswirtschaftslehre
D-76131 Karlsruhe

Dieser Bericht einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung der Herausgeber unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Osnabrück, März 2004

ISBN 3-936475-13-X

Inhaltsverzeichnis

INHALTSVERZEICHNIS.....	I
ABBILDUNGSVERZEICHNIS	III
TABELLENVERZEICHNIS.....	IV
1 AUSGANGSLAGE.....	1
2 LEITFADEN ZUR PLANUNG DES EVALUATIONSPROZESSES...3	3
3 ZIELE UND FUNKTIONEN DER EVALUATION	5
4 EVALUATIONSBEREICHE	8
4.1 LERNUMGEBUNG	9
4.1.1 LERNKULTUR	9
4.1.2 LERNANGEBOT.....	10
4.2 LERNERVORAUSSETZUNGEN.....	12
4.2.1 KOGNITIVE FAKTOREN	13
4.2.2 MOTIVATIONALE FAKTOREN	17
4.2.3 EMOTIONALE ASPEKTE	21
4.3 LERNERGEBNISSE	23
4.3.1 AKZEPTANZ	23
4.3.2 LERNERFOLG.....	26
4.3.3 LERNTRANSFER.....	28
5 EVALUATIONSMETHODEN.....	31
5.1 DATENERHEBUNGSMETHODEN	31
5.1.1 DOKUMENTENANALYSE.....	31
5.1.2 BEFRAGUNG.....	31
5.1.3 BEOBACHTUNG	37
5.1.4 VERHALTENSRECORDING.....	39
5.1.5 TEST.....	41

5.1.6 EXPERTENBEURTEILUNG.....	42
5.2 DATENAUSWERTUNG.....	46
5.2.1 QUALITATIVE AUSWERTUNGSVERFAHREN	47
5.2.2 QUANTITATIVE AUSWERTUNG	48
6 BERICHTLEGUNG.....	56
7 PROBLEME BEI DER EVALUATION	58
8 SCHLUSSBEMERKUNGEN	61
LITERATURVERZEICHNIS.....	63

Abbildungsverzeichnis

ABBILDUNG 1:	EVALUATIONSBEREICHE	8
ABBILDUNG 2:	BEOBACHTUNGSPROTOKOLL	38
ABBILDUNG 3:	BEOBACHTUNGSKONTROLLE IM ZEITVERLAUF	38
ABBILDUNG 4:	HIERARCHISCHE STRUKTUR VOM KRITERIENKATALOG MEDA	44
ABBILDUNG 5:	BEWERTUNGSBILDSCHIRM VON AKAB.....	45
ABBILDUNG 6:	ANWENDUNGSBEISPIEL DER STRUKTUR-LEGE-TECHNIK ...	47
ABBILDUNG 7:	GRAPHISCHE DARSTELLUNG DER MITTELWERTE	51
ABBILDUNG 8:	GRAPHISCHE DARSTELLUNG DER INTERAKTION EINER ZWEIFAKTORIELLEN VARIANZANALYSE	55

Tabellenverzeichnis

TABELLE 1:	ITEMAUSZUG ZUR BEWERTUNG DER INDIVIDUALISIERUNGSMÖGLICHKEITEN	11
TABELLE 2:	ITEMAUSZUG ZUR BEWERTUNG DER ANIMATION	12
TABELLE 3:	ITEMAUSZUG ZUR ERFASSUNG DER INFORMATIONSTRATEGIEN	15
TABELLE 4:	ITEMAUSZUG ZUR ERFASSUNG DER KONTROLL- STRATEGIEN	16
TABELLE 5:	ITEMAUSZUG ZUR ERFASSUNG DER STÜTZSTRATEGIEN	17
TABELLE 6:	ITEMAUSZUG ZUR ERFASSUNG DER MOTIVATION	18
TABELLE 7:	ITEMAUSZUG ZUR ERFASSUNG DES STELLENWERTES VON COMPUTERKENNTNISSEN	20
TABELLE 8:	ITEMAUSZUG ZUR ERFASSUNG DES STELLENWERTES VON E-LEARNING	20
TABELLE 9:	ITEMAUSZUG ZUR ERFASSUNG DER COMPUTERBEZOGENEN ERFOLGSERWARTUNG	21
TABELLE 10:	ITEMAUSZUG ZUR ERFASSUNG DES EMOTIONALEN BEFINDENS DER LERNER	23
TABELLE 11:	ITEMAUSZUG ZUR ERFASSUNG DER GENERELLEN AKZEPTANZ	25
TABELLE 12:	ITEMAUSZUG ZUR ERFASSUNG DER AKZEPTANZ EINES LERNANGEBOTS	26
TABELLE 13:	ITEMAUSZUG ZUR ERFASSUNG DES LERNTRANSFERS	30
TABELLE 14:	LEITFADENAUSZUG ZUR ERFASSUNG DER EINSTELLUNG ZUM COMPUTER/INTERNET	32
TABELLE 15:	BEISPIELE FÜR TRANSKRIPTIONSZEICHEN	33
TABELLE 16:	BEISPIELE MIT UNTERSCHIEDLICHEN ANTWORTMÖGLICH- KEITEN FÜR DIE SCHRIFTLICHE BEFRAGUNG	35
TABELLE 17:	BENUTZERPROTOKOLL	40
TABELLE 18:	ITEMBEISPIEL FÜR DIE INTERVALLSKALA	49
TABELLE 19:	TABELLARISCHE DARSTELLUNG DER HÄUFIGKEITSVERTEILUNG	50
TABELLE 20:	AUSZUG AUS EINER TABELLARISCHEN DARSTELLUNG VON MITTELWERTEN UND STANDARDABWEICHUNGEN	52

TABELLE 21:	AUSZUG AUS EINER TABELLARISCHEN DARSTELLUNG DER χ^2 -ANALYSE.....	53
TABELLE 22:	AUSZUG AUS EINER TABELLARISCHEN DARSTELLUNG DER ERGEBNISSE EINER ZWEIFAKTORIELLEN VARIANZANALYSE	54
TABELLE 23:	GRÜNDE FÜR MANGELNDE AKZEPTANZ DER EVALUATION (AUSZUG)	58
TABELLE 24:	GRUNDLAGE ZUR KLÄRUNG DER ZIELVORSTELLUNGEN	59

1 Ausgangslage

E-Learning, das computer- und netzbasierte Lernen, verspricht vielfältige Möglichkeiten für eine effektive Bildung und Qualifikation in Schulen, Hochschulen und Unternehmungen. Vor allem deshalb ist in den letzten Jahren auf allen Ebenen der Aus- und Weiterbildung eine deutliche Zunahme computer- und netzbasierter Lernangebote zu verzeichnen.

Die aktuelle Situation ist allerdings durch Unsicherheit und Intransparenz im Hinblick auf die *tatsächliche* pädagogische und fachliche Leistungsfähigkeit der Lernangebote gekennzeichnet. Eine systematische Evaluation der Arrangements und der Lehr-Lernprozesse findet bisher nur in unzureichendem Maße statt. Die Ursachen dafür sind vielfältig: Zum einen ist Evaluation ein aufwendiger, häufig kostenintensiver Prozess. Zum anderen gibt es kaum ganzheitliche Konzepte, die theoriegestützte und zugleich praxiserprobte Strategien, Methoden und Kriterien für die Evaluation der computer- und netzbasierten Lernangebote bereitstellen. Die vorliegende Publikation greift dieses Defizit auf. Anhand des BMBF-Projektes „Interdisziplinäres multimediales Programm für universitäre Lehre und selbstorganisiertes Lernen zum Thema Electronic Commerce“ (IMPULS^{EC})¹ wird ein Evaluationskonzept vorgestellt und werden das methodische Vorgehen sowie die Instrumente theoriegeleitet beschrieben und exemplarisch Ergebnisse präsentiert.

Evaluation ist ein vielschichtiger Begriff, der je nach seiner Zuordnung zu verschiedenen Bereichen ganz unterschiedlich akzentuiert werden kann. Verwandte Begriffe sind Begleitforschung, Erfolgskontrolle, Wirkungskontrolle oder Qualitätskontrolle. Im Kontext dieser Publikation und bezogen auf den in Rede stehenden Gegenstand (computer- und

¹ Das Projekt IMPULSEC (Interdisziplinäres multimediales Programm für universitäre Lehre und selbstorganisiertes Lernen zum Thema „Electronic Commerce“) erstellt im Rahmen des BMBF-Programms „Neue Medien in der Bildung“ einen multimedialen internetgestützten Lehrgang zum Thema „E-Commerce“, der in der Ausbildung von Studierenden wirtschaftswissenschaftlicher Studiengänge an den beteiligten Standorten und damit auch in der Handelslehrausbildung eingesetzt wird. Projektpartner sind neben dem Lehrstuhl für Berufs- und Wirtschaftspädagogik an der Universität Leipzig Wirtschaftsinformatiker und Betriebswirte der Universitäten in Dresden, Osnabrück, Würzburg und Karlsruhe. Der Lehrgang umfasst 11 Kurse (vgl. <http://impuls01.wiwi.tu-dresden.de/index.html>).

netzbasierte Lernangebote) wird unter Evaluation die systematische und zielgerichtete Sammlung, Analyse und Bewertung von Daten zur Qualitätssicherung und Qualitätskontrolle der Lernangebote verstanden.

Im Rahmen dieser allgemeinen Definition sind je nach spezifischen Zielsetzungen und Realisierungsmöglichkeiten unterschiedliche Evaluationskonzepte, sog. „Evaluationsmodelle“ (vgl. Wottawa, 2001) entwickelt worden. Wulf (1972; 1975) unterscheidet beispielsweise zwischen praxisorientierter, entwicklungsorientierter und theorieorientierter Evaluation. Besonders bedeutsam für das Lernen und Lehren mit computer- und netzbasierten Lernangeboten ist die auf Scriven (1991) zurückgehende Einteilung in *summative* und *formative* Evaluation. Unter der *summativen Evaluation* wird die Überprüfung der Resultate verstanden, nachdem eine Bildungsmaßnahme abgeschlossen worden ist. Die *formative Evaluation* bezieht sich auf die Durchführung der Lehr-Lern-Prozesse und deren Wirkung: Sie verfolgt das Ziel, die Qualität des Lernangebots fortlaufend zu kontrollieren und auf der Basis der gewonnenen Daten zu verbessern.

Diese Systematiken können zwar als allgemeine Orientierungsraster bei der Vorbereitung von Evaluationsstudien hilfreich sein. Das praktische Vorgehen bei der Evaluation konkreter Forschungs- und Entwicklungsprojekte ist jedoch zumeist außerordentlich komplex und entzieht sich letztlich einer vollständigen Beschreibung mithilfe abstrakter Klassifizierungen. Der vorliegende Text geht deshalb darüber hinaus. Er thematisiert die Art und Weise der Planung und Realisierung von Evaluationsstudien auf einer Ebene konkreter Handlungsschritte und mit hoher Anschaulichkeit.

2 Leitfragen zur Planung des Evaluationsprozesses

Die Planung der Evaluation ist ein komplexer Prozess, der zumindest die folgenden Fragen beantworten sollte:

- *Welche Ziele werden mit der Evaluation verfolgt?*

Es muss festgelegt werden, welchen Funktionen die Evaluation dienen soll: zum Beispiel der Schwachstellenermittlung zur Optimierung der Entwicklung eines Lehr-Lern-Arrangements oder der Entscheidung darüber, ob ein Lernprogramm gekauft und eingesetzt wird.

- *Was soll evaluiert werden?*

Es gibt unterschiedliche Evaluationsbereiche, die identifiziert und bearbeitet werden können bzw. sollten oder müssen. Für computer- und netzbasierte Lernangebote sind neben der Lernumgebung z. B. auch noch die Lernvoraussetzungen und Lernergebnisse zu berücksichtigen.

Unter der Lernumgebung wird nicht nur das Lernangebot selbst verstanden, dazu zählen u. a. auch die Rahmenbedingungen, wie zum Beispiel die Lernkultur. Die Lernergebnisse werden über den Lernerfolg und die Akzeptanz erfasst. Sie sind von den kognitiven, motivationalen und emotionalen Lernvoraussetzungen abhängig.

- *Welche Methoden werden verwendet?*

Es gibt verschiedene Evaluationsmethoden, die sowohl quantitative als auch qualitative Analysen ermöglichen. Dazu gehören u. a. Interviews, die Anwendung von Checklisten, Befragungen, Dokumentenanalysen und Beobachtungen. Alle Methoden haben Vor- und Nachteile in Bezug auf die Kosten und die Informationshaltigkeit der Daten. Die Wahl der Methoden erfolgt in Abhängigkeit von den Zielen und vom Zweck der Evaluation sowie vom Evaluationsdesign und den zur Verfügung stehenden Ressourcen.

- *Welche Evaluationsergebnisse entstehen und wie werden sie verwendet?*

Die Evaluationsergebnisse können z. B. intern für die Schwachstellenermittlung und Optimierung des Lernangebots genutzt werden. Sie können aber auch im Unternehmen veröffentlicht oder bewusst nach „außen“ publiziert werden.

- *Welche Probleme können auftreten?*

Bei der Evaluation können verschiedene Störungen auftreten. Dazu zählen z. B. Unruhe unter den Beteiligten durch mangelnde Akzeptanz und höhere Arbeitsbelastung oder Zeitverzögerungen bei der Entwicklung und beim Einsatz des Lernangebots. Solche Probleme sind bereits bei der Planung der Evaluation einzukalkulieren und durch entsprechende Maßnahmen abzufedern.

Die folgenden Erörterungen werden durch die dargestellten handlungsleitenden Fragestellungen zur Planung des Evaluationsprozesses strukturiert.

3 Ziele und Funktionen der Evaluation

Evaluation im Kontext von multimedialen Lernumgebungen kann sich sowohl auf einzelne Produktmerkmale als auch auf den gesamten Prozess der Planung, der Entwicklung und des Einsatzes beziehen. In diesem Zusammenhang ist zwischen der Produktevaluation und der Prozessevaluation zu unterscheiden. Bei der Evaluation einzelner Produktmerkmale von Bildungsangeboten werden Inhalte, Methoden, Medien und deren Zusammenwirken unter den Aspekten von Qualität, Funktionalität, Effektivität und Nutzen beurteilt. Für die Prozessevaluation sind vor allem die Planungs- und Entscheidungsprozesse bei der Entwicklung einer Lernumgebung ausschlaggebend.

In der Praxis treten je nach der Zielsetzung unterschiedliche Varianten der Funktionen von Evaluation auf, zum Beispiel abhängig davon, ob die Evaluationsergebnisse nach „innen“ oder nach „außen“ (u. a. an Bildungsträger oder potentielle Anwender) gerichtet sind, oder ob sie bestimmte Konsequenzen für weitere Entscheidungen haben.

Nach den Verwertungsinteressen der Evaluationsergebnisse können drei Funktionen von Evaluation unterschieden werden (Rowntree, 1992):

- Erkenntnisfunktion,
- Kontroll- und Entscheidungsfunktion,
- strategisch-politische Funktion.

Die *Erkenntnisfunktion* bezieht sich auf die Generierung von Erkenntnissen über Effekte eines Lernangebots, sowohl bezogen auf das einzelne Produkt als auch im Vergleich mit anderen Angeboten. Die Erkenntnisgewinnung dient vielfach wissenschaftlichen Fragestellungen und ist eher selten mit einer Entscheidungskonsequenz verbunden.

Die *Kontroll- und Entscheidungsfunktion* ist nach innen gerichtet. Sie dient der Identifizierung von Schwachstellen und deren Behebung, der Optimierung des Lernangebots, sowie der Bewertung von Begleitumständen bei der Entwicklung. Besondere Bedeutung hat diese Funktion für Unternehmen, die ihre Lernangebote selbst entwickeln bzw. gekaufte Angebote an ihre Bedürfnisse anpassen.

Bei der *strategisch-politischen* Funktion geht es darum, die Evaluationsergebnisse eines Lernangebots nach „außen“ zu tragen. Diese Funktion ist in der Regel darauf gerichtet, den Nutzen einer Bildungsmaßnahme resp. eines Lernangebots z. B. gegenüber dem Bildungsträger, der Geschäftsführung oder potentiellen Anwendern darzustellen.

Die Funktionen der Evaluation sind zudem in Abhängigkeit von der Projektphase unterschiedlich zu definieren, insbesondere, wenn ein E-Learning-Angebot selbst entwickelt wird. Dazu kann ein theoretisches Modell dienen, das die unterschiedlichen Zielsetzungen der Projektphasen berücksichtigt. Ein Beispiel dafür ist das Phasenmodell der Evaluation (vgl. Tergan, 2000), das im Folgenden dargestellt wird. Es geht davon aus, dass die Qualitätssicherung nicht nur die Qualität des Designs oder die Bewertung der Wirkung eines Lernangebots sondern auch die Analyse der Rahmenbedingungen der Bildungsmaßnahmen betrifft. Das Phasenmodell sieht drei Phasen vor: *Planungs-, Entwicklungs- und Einsatzphase*.

Ziel der *Planungsphase* ist es, mit einer Zielgruppenanalyse Daten zu gewinnen, die für die Optimierung der Projektkonzeption genutzt werden können. Außer der Ermittlung des Bildungsbedarfs und der Bewertung der Ziele und Rahmenbedingungen liefert die Zielgruppenanalyse wichtige Informationen über Lernervoraussetzungen und über die Erfahrungen, Einstellungen und Erwartungen der Zielgruppe. Diese Informationen können zur Formulierung bzw. Korrektur der Projektkonzeption beitragen. Ein Beispiel derartiger empirischer Studien ist die Zielgruppenanalyse von Klauser & Kim (2003a).

In der *Entwicklungsphase* wird die Qualität des in der Entwicklung befindlichen Lernangebots mehrfach überprüft. Das Evaluationsziel ist die Sicherung und Optimierung der pädagogischen Qualität. Evaluiert wird mit der Fragestellung, ob das eingesetzte Lernangebot – also Methoden und Medien - den zu vermittelnden Inhalten und der Gesamtkonzeption entspricht.

Die Evaluation in der *Einsatzphase* hat in der Regel sowohl formative als auch summative Anteile. Während die summative Evaluation auf die Wirkungen gerichtet ist, dient die formative Evaluation in dieser Phase der

Gewinnung praxisbezogener Daten. Mit deren Hilfe soll das Lernangebot weiter verbessert werden. Für die formative Evaluation steht die Kontroll- und Entscheidungsfunktion im Vordergrund (Rowntree, 1992). Sie dient der Optimierung, der Identifizierung von Schwachstellen und deren Behebung sowie der Bewertung von Begleitumständen der Entwicklung. Evaluationsgegenstände sind hier also Wirkungen und Nutzen des Angebots bzw. seiner einzelnen Komponenten. Zudem dient eine wissenschaftlich konzipierte Evaluation auch der Überprüfung der in der Planungsphase präzisierten Zielsetzungen. Wichtige Kriterien sind dabei u. a. Akzeptanz, Lerneffekte und bildungspolitische, organisatorische sowie ökonomische Wirkungen.

Entwicklung und Einsatz eines E-Learning-Angebots ist ein komplexer Prozess, dessen Erfolg von vielen Faktoren abhängig ist, die einander wechselseitig beeinflussen. Darauf wird im Folgenden eingegangen. Es wird gezeigt, welche Evaluationsbereiche zu erfassen sind, um die Qualität eines computer- und netzbasierten Lernangebots systematisch zu sichern und zu kontrollieren.

4 Evaluationsbereiche

In diesem Abschnitt werden die unterschiedlichen Evaluationsbereiche vorgestellt. Dies geschieht anhand eines Modells, das sich für multimediale Lernangebote als zweckmäßig erwiesen hat. Es werden drei Bereiche bzw. Evaluationsgegenstände unterschieden: Lernumgebung, Lernervoraussetzungen und Lernergebnisse (siehe Abbildung 1).

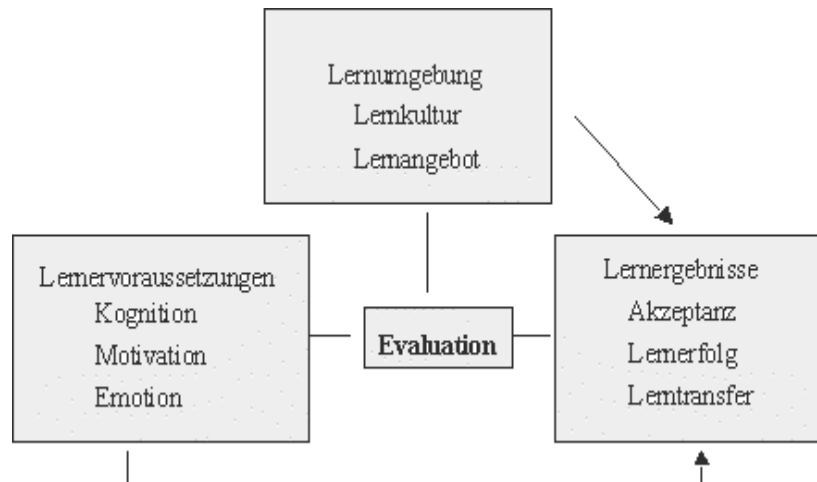


Abbildung 1: Evaluationsbereiche

Das Ziel jeder Evaluation ist es, die Zusammenhänge zwischen den einzelnen Evaluationsbereichen zu analysieren und Aussagen zur Qualitätssicherung und Qualitätskontrolle der Lehr-Lernprozesse zu treffen.

Die Lernergebnisse, die über die Evaluationsbereiche „Lernerfolg“ (Leistungs- bzw. Verhaltensänderung) und „Akzeptanz“ erfasst werden, sind das eigentliche Evaluationsobjekt. Sie können aber nur in Abhängigkeit von den beiden anderen Bereichen, der Lernumgebung und den Lernervoraussetzungen, adäquat interpretiert werden.

Unter Lernumgebung wird nicht nur das Lernangebot selbst verstanden, sondern dazu zählen auch die Rahmenbedingungen, wie zum Beispiel die Lernkultur, von welcher der Erfolg der Implementation abhängig ist. Lerner-

voraussetzungen, die als kognitive, motivationale und emotionale Merkmale der Lernenden erfasst werden, sind ebenfalls wichtige Analysegegenstände der Evaluation (vgl. Dillon & Gabbard, 1998; Klauser & Kim, 2003a).

4.1 Lernumgebung

Die Lernumgebung ist nicht nur als physikalische Umgebung, sondern im Sinne eines Arrangements zu interpretieren. Sie umfasst a) unterschiedliche Materialien, Informationsträger und Hilfen, die aufeinander abgestimmt sind, um Lernprozesse anzuregen und zu unterstützen - das Lernangebot - und b) vor allem eine spezifische Lernkultur.

4.1.1 Lernkultur

E-Learning stellt neue Potenziale bereit, um den veränderten und erweiterten Bildungsbedarf der Informations- und Wissensgesellschaft zu befriedigen. Eine wichtige Voraussetzung für die Etablierung und die fortschreitende Verbreitung der spezifischen Formen computer- und netz-basierten Lernens ist ein Grundkonsens der Beteiligten im Hinblick auf die Ziele und ein effektives Handeln im Kontext der neuen Medien. In diesem Zusammenhang zeigt sich die Bedeutsamkeit der Lernkultur als Grundvoraussetzung der Förderung von Potenzialen.

Lernkultur in diesem Sinne umfasst vor allem

- das Bewusstsein über den Stellenwert des Lernens,
- die Einstellung bzw. affektiv-kognitive Bereitschaft der Beteiligten, sich auf Lernprozesse einzulassen und sich für die neuen Erfahrungen zu öffnen, und
- die Unterstützung, mit der eine lernfähige und potenzialorientierte Kultur überhaupt erst betrieben werden kann (Sonntag, 1996).

Bei der Evaluation der Lernkultur können neben Informationen über die normativen Setzungen der Unternehmenspolitik oder Leitbilder des Unternehmens weitere Punkte durch Befragung erfasst werden, zum Beispiel:

- inwieweit Ressourcen organisationaler Art wie Räume oder Ausstattungen und personaler Art wie Einstellungen oder

Lernbereitschaft der Mitarbeiter und Führungskräfte dem Bildungsmanagement zur Verfügung stehen,

- inwieweit Lernen und Potenzialförderung als integraler Bestandteil der Unternehmensplanung angesehen werden,
- inwieweit die Mitarbeiterschulung in die betriebliche Karriereplanung eingebunden werden kann,
- inwieweit der Arbeitsplatz mit Lernmöglichkeiten verbunden werden kann,
- ob den Mitarbeitern die benötigte Zeit zur Verfügung gestellt wird,
- ob Belohnungen für die Durchführung und Beendigung eines Kurses festgesetzt werden,
- wie transparent der Lernbedarf und die Lernziele ermittelt worden sind,
- ob die Beteiligten über den aktuellen Stand der Bildungsmaßnahmen laufend informiert werden und
- wie groß das Ausmaß der Partizipation aller Organisationsmitglieder am Lernprozess ist (Sonntag, 1996).

4.1.2 Lernangebot

Das Lernangebot ist der Kern der Lernumgebung, auf das sich die Lernprozesse im engeren Sinne beziehen. Die Bewertung von Lernangeboten erfolgt vielfach mit Hilfe von Kriterienkatalogen. Obwohl die Evaluation mit Kriterienkatalogen zumeist als „Expertenevaluation“ verstanden wird, ist es wünschenswert, dass sowohl Experten als auch Lerner an der Beurteilung des Lernangebots beteiligt werden. In diesem Fall ist es jedoch notwendig, die Anzahl und die Formulierung der Kriterien dementsprechend anzupassen.

Kriterienkataloge zur Evaluation von Lernangeboten sollten ganzheitlich gestaltet sein, d.h., die Kataloge sollten medienpädagogische, technische und ökonomische Aspekte umfassen². Da E-Learning zumeist auf selbst-

2 Ein solcher ganzheitlich konzipierter Kriterienkatalog wurde für das Forschungsprojekt eLBe entwickelt, das im Auftrag des Sächsischen Staatsministeriums für Kultus am Lehrstuhl für Berufs- und Wirtschaftspädagogik der Universität Leipzig bearbeitet wurde (vgl. Sächsisches Staatsministerium für Kultus, 2002; Klauser, Kim & Pollmer, 2003).

gesteuertes Lernen abzielt, sind solche medienpädagogischen Kriterien, wie Lernzielformulierung, Orientierung, Individualisierungsmöglichkeiten, Lernerunterstützung und Lernerfolgskontrolle von besonderer Relevanz. Die Evaluation von Individualisierungsmöglichkeiten kann beispielsweise mit folgenden Items realisiert werden:

Kann der Nutzer zwischen verschiedenen Schwierigkeitsstufen wählen?	Ja <input type="radio"/>	Nein <input type="radio"/>
Kann der Nutzer das Lerntempo selbst bestimmen?	Ja <input type="radio"/>	Nein <input type="radio"/>
Kann der Nutzer einen individuellen Lernweg wählen?	Ja <input type="radio"/>	Nein <input type="radio"/>
Kann der Nutzer seine Arbeitsergebnisse individuell speichern?	Ja <input type="radio"/>	Nein <input type="radio"/>

Tabelle 1: Itemsauszug zur Bewertung der Individualisierungsmöglichkeiten

Hinzu kommen Aspekte, die die multimediale Gestaltung betreffen, wie Grafik, Animation und akustische Gestaltung. Um die pädagogischen Funktionen zu unterstützen, sollten auch technische Kriterien wie optimale Ladezeit, Login/Logout und Dateneingabe/-ausgabe/-speicherung/-schutz gesichert werden.

Folgendes Beispiel zeigt einen Itemsauszug zur Bewertung von Animationen:

Sind Animationen vorhanden?
Ja <input type="radio"/> Nein <input type="radio"/>
Wenn ja, dienen diese der Verdeutlichung von Lerninhalten?
wenig verdeutlichend <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> sehr verdeutlichend
Sind die Bewegungsabläufe gleichmäßig?
Ja <input type="radio"/> Nein <input type="radio"/>
Kann der Lernende die Darstellungsgeschwindigkeit beeinflussen?
Ja <input type="radio"/> Nein <input type="radio"/>

Tabelle 2: Itemsauszug zur Bewertung der Animation

Immer größere Bedeutung gewinnen ökonomische Aspekte wie das Kosten-Leistungs-Verhältnis eines Lernangebots. Die Zusammensetzung der Bewertungskriterien erfolgt dabei in Abhängigkeit von der Zielsetzung der Bildungsmaßnahme.

Um die Qualität eines E-Learning-Angebots umfassend zu beurteilen, müssen sowohl medienpädagogische und technische als auch ökonomische Betrachtungsaspekte zu einer integrativen Sichtweise zusammengeführt werden. Die ausgewogene Berücksichtigung aller drei Komponenten in ihrer Gesamtheit gilt als eine entscheidende Bedingung dafür, dass die E-Learning-Angebote und die Lehr-Lern-Prozesse sowohl der Zielgruppe als auch den pädagogischen Zielsetzungen sowie den konkreten Einsatzbedingungen gerecht werden können.

4.2 Lernervoraussetzungen

Im Kontext der Aus- und Weiterbildung wird häufig die Bedeutung des „maßgeschneiderten Lernens“ betont (Mandl & Reinmann-Rothmeier, 2000): Das Lernangebot soll den individuellen Bedürfnissen der einzelnen Lernenden angepasst werden bzw. angepasst sein. In diesem Zusammenhang wird den Lernenden eine zentrale Rolle bei der Sicherung des Lern- bzw. Transfererfolgs zugeschrieben: In Lernumgebungen, in denen selbstgesteuertes Lernen im Vordergrund steht, wird ein Lernender nicht als

passiver „Adressat“ sondern als aktiv Handelnder verstanden. Er setzt sich mit dem Lernangebot auseinander und gestaltet den Lernprozess weitgehend eigenverantwortlich. Lernervoraussetzungen, wie Interesse, Vorerfahrungen oder Motivation gelten in diesem Kontext als Determinanten für Erfolg und Effizienz. Wie in der pädagogischen Psychologie üblich, wird dabei zwischen kognitiven, motivationalen und emotionalen Faktoren unterschieden.

4.2.1 Kognitive Faktoren

Als kognitive Lernervoraussetzungen werden außer dem Vorwissen die Lernstrategien erfasst, die direkte Einflüsse auf den Lernprozess haben.

Vorwissen

Vorwissen ist mit einem „Eingangstest“ oder durch eine Lernerfolgskontrolle der vorangegangenen Lerneinheiten zu erfassen. Die wichtigste Methode der Vorwissensfeststellung ist der Test, dessen Ziel in der Erfassung der Voraussetzung für den Lernerfolg liegt. Zunächst muss das notwendige Vorwissen für den Lernerfolg identifiziert werden. Daraus können die Items zur Erfassung von vorhandenem Vorwissen abgeleitet werden, welches als erreichter Punktwert im Test ermittelt wird.

Das notwendige und das vorhandene Vorwissen können nur bezogen auf den Lerninhalt erfasst werden. Daher ist es sinnvoll, dass die Items zur Erfassung von Vorwissen durch den Autor des Lerninhalts erstellt werden. Außerdem sollte ein Kriterium festgelegt werden, um zu entscheiden, ob einem Lernenden der Zugang zum Lernangebot erteilt wird, und wie ein fortgeschrittener Lerner eingestuft wird. Je mehr ein Lernangebot den Lernenden Freiräume anbietet, angepasst an ihr Vorwissensniveau, zu lernen, desto mehr kann von der Individualisierbarkeit des Lernangebots die Rede sein. Außerdem können die Ergebnisse des Vorwissenstests als zusätzliche Information für die Interpretation der Lernerfolgskontrolle dienen.

Lernstrategien

Lernstrategien werden als mental repräsentierte Schemata oder Handlungspläne zur Steuerung des eigenen Lernverhaltens definiert. Sie setzen sich aus einzelnen Handlungssequenzen zusammen und sind

situationsspezifisch abrufbar. Durch die Nutzung solcher Strategien können Gedächtnisleistungen verbessert und Wissensstrukturen effektiv organisiert werden. Lernstrategien haben habituellen Charakter, in dem Sinne, dass jeder Mensch Präferenzen für gewisse Strategien entwickelt und diese dann bevorzugt einsetzt. Dabei ist jedoch zwischen Kompetenz und Performanz zu unterscheiden, denn das Wissen über bestimmte Lernstrategien bedeutet noch nicht, dass sie vom Lerner auch regelmäßig eingesetzt werden. Aufgabe der Evaluation ist es, die tatsächlich angewendeten Lernstrategien zu erfassen.

Das häufig angewendete Klassifikationsmodell unterscheidet folgende drei Arten der Lernstrategien (Wild, 2000):

- Informationsstrategien,
- Kontrollstrategien,
- Stützstrategien.

Informationsstrategien werden auch als kognitive Strategien bezeichnet und dienen der „unmittelbaren Informationsaufnahme, Informationsverarbeitung und -speicherung“. Sie umfassen Vorgänge wie *Organisieren*, *Elaborieren*, *Wiederholen* und *kritisches Prüfen*.

- *Organisieren* bezieht sich auf die Lerntätigkeit, den zu bewältigenden Lernstoff in geeigneter Weise zu reorganisieren. Es umfasst das Reduzieren und Ordnen von neuen Informationen beispielsweise durch Zusammenfassung, Gliederung und Anfertigung von Tabellen bzw. Schaubildern.
- *Elaborieren* meint das Integrieren neuer Informationen in die vorhandene Wissensstruktur, zum Beispiel durch Querverbindungen zu anderen Sachverhalten.
- *Wiederholen* dient der Verankerung im Gedächtnis. Es bezieht sich auf das Einprägen von Lernstoffen durch schlichtes Wiederholen.
- *Kritisches Prüfen* meint das Hinterfragen und Nachdenken über Sachverhalte. Es vertieft das Verständnis.

Lernstrategien	Beispiele
Organisieren	„Ich mache mir kurze Zusammenfassungen der wichtigsten Inhalte als Gedankenstütze.“
Elaborieren	„Ich versuche, neue Konzepte auf mir bereits bekannte Begriffe und Theorien zu beziehen.“
Wiederholen	„Ich lerne den Lernstoff anhand von Skripten oder anderen Aufzeichnungen möglichst auswendig.“
Kritisches Prüfen	„Ich frage mich, ob der Text, den ich gerade durcharbeite, wirklich überzeugend ist.“

Tabelle 3: Itemsauszug zur Erfassung der Informationsstrategien

Kontrollstrategien sind interne Kontrollvorgänge. Sie werden deshalb auch als metakognitive Strategien bezeichnet. Sie regeln den selbstgesteuerten Lernprozess und umfassen Vorgänge wie Planen, Überwachen und Regulieren.

- *Planen* bezieht sich auf die inhaltliche Vorbereitung des Lernvorgangs.
- *Überwachen* umfasst Aktivitäten zur Kontrolle, die als Ist-Soll-Vergleich durchgeführt werden.
- *Regulieren* meint Verhaltensänderungen, wenn es die Situation erfordert, also wenn man eigene Lernschwierigkeiten bei einem Thema feststellt.

Lernstrategien	Beispiele
Planen	„Vor dem Lesen überlege ich mir, welche Fragen mir der Text beantworten soll.“
Überwachen	„Ich stelle mir Fragen zum Stoff, um sicherzugehen, dass ich auch alles verstanden habe.“
Regulieren	“Wenn ich ein schwieriges Themengebiet vorliegen habe, passe ich meine Lerntechnik den höheren Anforderungen an.“

Tabelle 4: Itemsauszug zur Erfassung der Kontrollstrategien

Im Gegensatz zu den beiden vorher genannten Kategorien handelt es sich bei den *Stützstrategien* um Sekundär- bzw. ressourcenbezogene Strategien. Sie sind auf Ressourcen bezogen, auf die der Lerner zurückgreifen kann, um den Lernerfolg zu optimieren. Sie umfassen *Zeitmanagement*, *die Gestaltung der Lernumgebung und des Arbeitsplatzes*, *die Nutzung zusätzlicher Informationsquellen* und *das Lernen mit Kollegen*.

- *Zeitmanagement* meint, dass E-Learning die Möglichkeit bietet, von Zeit und Ort unabhängig zu lernen. Für effektives Lernen am Arbeitsplatz sind jedoch die Festlegung und das Einhalten eines detaillierten Zeitplans von großer Bedeutung.
- *Die Gestaltung der Lernumgebung* verweist darauf, dass eine positive Lernumgebung den Lernprozess unterstützen kann. Deshalb sind die Wahl des Lernplatzes und ein gut organisierter Arbeitsplatz sehr wichtig.
- *Die Nutzung zusätzlicher Informationsquellen* ist ein weiterer wichtiger Punkt. Beim selbstgesteuerten Lernen kommt dem richtigen Umgang mit zusätzlichen Informationsquellen eine enorme Bedeutung zu.
- Beim *Lernen mit Peers* kann man sich über individuelle Sichtweisen unterhalten, sich den Stoff gegenseitig erklären - also Wissenslücken erkennen und schließen - und durch das Erklären das Stoffgebiet in eigenen Worten zusammenfassen.

Lernstrategien	Beispiele
Zeitmanagement	„Zum Zweck des effektiven Lernens lege ich einen Zeitplan fest.“
Gestaltung der Lernumgebung	„Während der Lernphase organisiere ich meinen Arbeitsplatz, so dass ich mich gut auf den Lernstoff konzentrieren kann.“
Nutzung zusätzlicher Informationsquellen	„Ich ziehe zusätzliche Informationsquellen heran, wenn meine Aufzeichnungen unvollständig sind.“
Lernen mit Peers	„Ich nehme die Hilfe anderer in Anspruch, wenn ich Verständnisprobleme habe.“

Tabelle 5: Itemsauszug zur Erfassung der Stützstrategien

4.2.2 Motivationale Faktoren

Lernmotivation

Die Motivationsforschung beschäftigt sich mit der Frage nach den Gründen und Absichten menschlichen Verhaltens. Bezogen auf Lernhandlungen kann die Lernmotivation als die Absicht oder die Bereitschaft einer Person definiert werden, sich ausdauernd mit einem Lerngegenstand auseinander zu setzen.

Um das komplexe Konstrukt der Motivation zu erfassen, können vier verschiedene Motivationsaspekte unterschieden werden (vgl. Prenzel & Drechsel, 1996):

- Intrinsische Motivation,
- Zukunftsorientierte Motivation,
- Rollenbezogene Motivation,
- Externale Motivation.

Die *intrinsische Motivation* ist dadurch gekennzeichnet, dass sie aus der Handlung selbst erwächst, also keiner Verstärkung von außen bedarf, da sich die Handlung von selbst verstärkt. Die Handlung selbst ist also der Zweck.

Eine Person mit *zukunftsorientierter Motivation* erachtet ihr Lernen als wichtig, um eigene Ziele erreichen zu können. Die Person handelt in Übereinstimmung mit ihren eigenen Werten und Zielen.

Personen mit *rollenbezogener Motivation* weisen häufig darauf hin, dass sie sich angestrengt haben, weil es von Lernenden erwartet wird. Die „rollenbezogene Motivation“ gibt Hinweise darauf, inwieweit die Teilnehmer sich entsprechend ihrer Rolle engagiert haben.

Bei *externaler Motivation* geht es nicht um die Lerninhalte sondern nur darum, die Ansprüche einer externen Instanz zu erfüllen. „Externale Motivation“ liegt dann vor, wenn die Teilnehmer deshalb lernen, weil sie Bestrafungen vermeiden oder Belohnungen erlangen wollen. Folgendes Beispiel zeigt den Itemsauszug zur Erfassung der Motivation.

Motivation	Beispiele
Intrinsische Motivation	„Beim Lernen hat mich die Sache so fasziniert, dass ich mich voll eingesetzt habe.“
Zukunftsbezogene Motivation	„Beim Lernen war mir klar, dass ich das für meinen Beruf können muss.“
Rollenbezogene Motivation	„Beim Lernen habe ich mich so angestrengt, wie es von Mitarbeitern erwartet wird.“
Externale Motivation	„Beim Lernen habe ich mich nur angestrengt, damit ich keinen Ärger bekomme.“

Tabelle 6: Itemsauszug zur Erfassung der Motivation

Stellenwert und Erfolgserwartung

Auf die Frage, wie es in einer konkreten Situation zu einer bestimmten Intentionsbildung kommt, oder warum eine spezifische Handlungsalternative

bevorzugt wird, gibt das handlungstheoretische Modell der Motivation eine Antwort (Heckhausen, 1989; Rheinberg, 1997). Dieses Modell geht davon aus, dass eine Person in der gegebenen Situation über mehrere Handlungsalternativen verfügt, von denen jede bestimmte Folgen nach sich zieht. In formaler Betrachtung errechnet sich der Grad der Handlungsbereitschaft - also der Motivation - aus dem Produkt der Erfolgserwartung und dem Wert der Handlungsfolgen. Für die Motivationsbildung sind also einerseits der Stellenwert der Handlung und andererseits die Erfolgserwartung wichtig. Wenn eine Aufgabe mit positiven Folgen verbunden ist, hat sie einen hohen Stellenwert. Wenn sie außerdem mit hoher Wahrscheinlichkeit erfolgreich vollzogen wird, neigt die Person noch stärker dazu, diese Aufgabe auszuführen. Die Erfolgserwartung hat Einfluss auf das kognitiv-motivationale Geschehen beim Lernen. Die Auswahl und Realisierung einer Handlungsalternative hängen schließlich davon ab, ob sie im Vergleich zu einer anderen Alternative eine positivere „Nutzen-Kosten-Bilanz“ erwarten lässt. Der Stellenwert von Computerkenntnissen und E-Learning kann zum Beispiel mit folgenden Items erfasst werden:

Die folgenden Aussagen beziehen sich auf den Stellenwert Ihrer Computerkenntnisse für Ihre Berufstätigkeit. Bitte geben Sie jeweils an, in welchem Ausmaß Sie den Aussagen zustimmen.

stimme zu stimme eher zu stimme eher nicht zu stimme nicht zu
 1 2 3 4

	1	2	3	4
Computerkenntnisse sind heute ein absolutes Muss für meine Berufstätigkeit.				
Wenn ich ausreichende Computerkenntnisse habe, wird es mir die Arbeit erleichtern.				
Für mich ist der Computer ein nützliches Arbeitsmittel.				

Tabelle 7: Itemsauszug zur Erfassung des Stellenwerts von Computerkenntnissen

Welchen Stellenwert hat E-Learning in Ihrer Lebenswelt? Bitte geben Sie an, wie wichtig E-Learning für Ihr Berufsleben ist.

sehr wichtig eher wichtig eher nicht wichtig nicht wichtig
 1 2 3 4

	1	2	3	4
Berufsleben - zur Zeit				
Berufsleben - in Zukunft				

Tabelle 8: Itemsauszug zur Erfassung des Stellenwerts von E-Learning

periodisch zu Beurteilungen der aktuellen Lage kommt. Je nachdem, ob die Einschätzungen im Hinblick auf die zu erreichenden Ziele positiv oder negativ ausfallen, kommt es zur Auslösung bestimmter Emotionen.

Emotionen können als „rudimentäre Motivationssysteme“ verstanden werden, die über Lust-Unlust-Qualitäten im Tätigkeitsvollzug die schon bestehende Lernmotivation steigern oder dämpfen. Angenehme Empfindungen lösen ein Bleiben bzw. Aufsuchen der Situation aus, unangenehme Befindlichkeit bewirkt das Gegenteil. So wird der Teilnehmer eines E-Learning-Kurses, der sich mit jeder weiteren nicht-verstandenen Lerneinheit wertloser und inkompetenter fühlt, eine starke Tendenz entwickeln, diesem Kurs zu entkommen. Dieses Verhalten wirkt unter Umständen der ursprünglichen Lernmotivation diametral entgegen.

Inwieweit sich Emotionen auf den Lern- und Leistungsprozess auswirken, hängt stark von den jeweiligen Aufgabenanforderungen ab. Entscheidend dabei ist, (1) ob die Aufgabe als *Herausforderung* angesehen wird oder nicht, und (2) wie groß die *Erfolgszuversicht* des Ausführenden ist. So wirkt sich die erlebte Herausforderung in Kombination mit hoher Erfolgszuversicht positiv auf die Leistung aus. Wird die Aufgabe als nicht herausfordernd erlebt und wird Misserfolg befürchtet, hat das negative Konsequenzen für den Lernprozess.

Bei der Erforschung des Zusammenhangs zwischen Emotionen und Leistung wurde festgestellt, dass der gesamte Stil der Informationsverarbeitung bereits durch leichte Stimmungsveränderungen beeinflusst wird. Zugleich findet man bei positiver Stimmung oft kreativere Lösungen und ein effektiveres Problemlöse- und Entscheidungsverhalten als bei negativer Stimmung.

Pekrun und Hofmann (1999) unterscheiden *drei Emotionskategorien*: (1) positive Emotionen wie Lernfreude, Leistungshoffnung, (2) aktivierende negative Emotionen, die auch physiologische und motivationale Effekte ausüben wie Angst oder Ärger und (3) deaktivierende Emotionen wie Langeweile und Hoffnungslosigkeit. Es wird angenommen, dass die Leistungswirkungen dieser Emotionen durch kognitive und motivationale Mediationsprozesse vermittelt werden.

Zur Erfassung des emotionalen Befindens der Lerner können unterschiedliche Emotionskategorien mit mehrstufigen Skalen vorgelegt werden.

Ich fühlte mich während der Online-Phase eher...				
trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	
1	2	3	4	
	1	2	3	4
a) frustriert				
b) herausgefordert				
c) gelangweilt				
d) interessiert				

(Quelle Kim & Klauser, 2004)

Tabelle 10: Itemsauszug zur Erfassung des emotionalen Befindens der Lerner

4.3 Lernergebnisse

4.3.1 Akzeptanz

In den in Rede stehenden Lernumgebungen resp. Lernangeboten steht das selbstgesteuerte Lernen im Vordergrund. Der Lernerfolg hängt davon ab, in welcher Art und Weise sich die Lernenden mit den vorgefundenen Bedingungen auseinandersetzen und ihren individuellen Lernprozess ausgestalten. Daher ist die Akzeptanz eines Lernangebots einer der wichtigsten Indikatoren für den Lernerfolg. Sie wird als die subjektive Wahrnehmung und Beurteilung der Qualität und des Nutzens eines Lernangebots durch Lerner definiert. Im Rahmen der Qualitätssicherung, insbesondere in ihrer formativen Evaluationsphase, nimmt die Analyse der Akzeptanz deshalb eine zentrale Stellung ein.

In der Akzeptanzforschung hat sich die Unterscheidung zwischen Einstellungsakzeptanz und Handlungsakzeptanz etabliert (Simon, 2001). Die Einstellungsakzeptanz fasst die affektive und kognitive Komponente zusammen. Sie umfasst sowohl die motivational-emotionale Haltung der Lerner als auch deren Kosten- und Nutzenabwägungen bei der Teilnahme an einer Bildungsmaßnahme. Der Nutzen ergibt sich aus dem Grad der Bedürfnisbefriedigung. Unter den „Nutzungskosten“ werden alle finanziellen Aufwendungen, Zeit und Anstrengungen verstanden. Von Handlungsakzeptanz wird gesprochen, wenn ein Schulungsangebot nicht nur positiv bewertet sondern auch in Form eines beobachtbaren Verhaltens (zum Beispiel Nutzung) angenommen wird. Im Hinblick auf die Akzeptanz eines Lernangebots können unterschiedliche Niveaus zwischen „Ablehnung“ und „Zustimmung“ identifiziert werden.

Die Akzeptanz kann generell oder für spezifische Beurteilungsaspekte erfragt werden. Der folgende Ausschnitt aus einem Fragebogen zeigt Beispiele zur Erfassung der generellen Akzeptanz.

Die folgenden Fragen beziehen sich darauf, wie Sie das E-Learning-Angebot insgesamt einschätzen. Bitte geben Sie jeweils an, in welchem Ausmaß Sie den Aussagen zustimmen.

trifft zu trifft eher zu trifft eher nicht zu trifft nicht zu
 1 2 3 4

	1	2	3	4
Das Lernen mit dem E-Learning-Angebot war effizient.				
Das Lernen mit dem E-Learning-Angebot war zu zeitaufwändig.				
Ich würde wieder mit einem E-Learning-Angebot lernen.				
Ich würde anderen Mitarbeitern das E-Learning-Angebot empfehlen.				

(Quelle Kim & Klauser, 2004)

Tabelle 11: Itemsauszug zur Erfassung der generellen Akzeptanz

Bei der differenziellen Akzeptanzanalyse können unterschiedliche Aspekte wie Inhalt und Design beurteilt werden. Es geht darum, inwieweit die Lernenden mit den einzelnen Aspekten des Lernangebots zufrieden sind. Die folgenden Tabellen zeigen Beispiele.

Lerninhalte	sehr zufrieden	sehr unzufrieden		
1. Interessant				
2. Verständlich				
3. Lernzielrelevant				
4. Praxisorientiert				

Design: Farbe 1. Anzahl 2. Farbkombination	sehr zufrieden					sehr unzufrieden
Design: Grafik 1. Häufigkeit 2. Größe 3. Position 4. Verständlichkeit	sehr zufrieden					sehr unzufrieden

(Quelle Kim & Klauser, 2004)

Tabelle 12: Itemsauszug zur Erfassung der Akzeptanz eines Lernangebots

4.3.2 Lernerfolg

Nach dem Einsatz eines Lernangebots hat die Evaluation eine Kontrollfunktion zu erfüllen. Dabei steht die Erfassung von Wirkungen wie Akzeptanz und Lernerfolg im Vordergrund. Die Evaluation des Lernerfolgs gibt dem Trainer bzw. Anbieter ein Feedback und kann somit zur Qualitätssicherung des Angebots beitragen. Das Feedback ermöglicht den Lernenden zudem eine individuelle Leistungseinschätzung, die sich wiederum motivationsfördernd auf den weiteren Lernprozess auswirken kann.

Lernerfolg als positives Lernergebnis ist durch die Einbeziehung von Lernzielen zu erfassen und zu erklären. Der Lernerfolg umfasst nicht nur Zuwachs an Wissen und Können auf der kognitiven Ebene, sondern auch die Veränderung auf der affektiven Ebene. Dazu zählen Einstellungs- und Motivationsänderungen. Wenn motorische Lernziele formuliert worden sind, ist zudem die Verbesserung dieser Fertigkeiten als Lernerfolg zu analysieren.

Häufig angewendete Methoden zur Feststellung des Lernerfolgs sind *Tests*.

Darunter versteht man Verfahren zur Messung der Ausprägung empirisch abgrenzbarer Verhaltens- und Leistungsmerkmale. Tests beziehen sich inhaltlich auf die im Lernangebot behandelten Themen. Das Ziel ist die

Erfassung der Wirkungen des Lernangebots auf den Lernerfolg, zum Beispiel gemessen als erreichter Punktwert in einem Behaltenstest nach der erfolgreichen Bearbeitung des Lernangebots oder einzelner Teile davon.

In der Praxis der Qualitätsevaluation werden häufig *informelle Tests* zur Lernerfolgserfassung verwendet. Sie sind nicht standardisiert oder kaum hinsichtlich der Gütekriterien, wie Objektivität, Reliabilität und Validität, überprüft worden. Diese Tests werden von Trainern oder Evaluatoren für wiederkehrende gleichartige Prüfungen nach Maßgabe der jeweiligen Evaluationsziele entwickelt und eingesetzt.

Kriteriumsorientierte Tests werden entwickelt, um eine inhaltlich differenzierte Leistungsbeurteilung zu gewährleisten. Die individuelle Leistung eines Lerners wird dabei nicht mehr an der Durchschnittsleistung der Mitlerner, sondern an einem sachlichen Kriterium, z. B. einem vorher festgelegten Lernziel, gemessen.

Es gibt Lerninhalte, für die sich *mündliche Prüfungen* als zweckmäßig erwiesen haben. Dabei wird die Lernleistung des Einzelnen auf der Grundlage eines Gesprächs beurteilt.

Aufsatzprüfungen ermöglichen es, Antworten frei zu formulieren und schriftlich niederzulegen.

Derartige Prüfungen sind jedoch, mehr als andere Methoden, von der fachlichen Kompetenz des Beurteilers abhängig. In der Praxis haben sich aus diesem Grunde Punktesysteme etabliert, die einzelne Aussagen und Argumentationsstrukturen nach einem festgelegten Schema mit einem Anforderungsprofil – also einer „Musterlösung“ – in Beziehung setzen. Dabei ist darauf zu achten, dass die maximal zu erreichende Punktzahl sowie die Gewichtung der einzelnen Aufgaben den Prüflingen mitgeteilt werden, um den Prozess der Leistungsermittlung und -bewertung transparent zu gestalten.

Eine weitere Form zur Lernergebniskontrolle insbesondere der affektiven Dimension ist die Befragung mit Fragebögen. Ein Fragebogen im Kontext computer- und netzbasierter Lernangebote kann z. B. die Änderung des Stellenwertes von E-Learning und die Veränderung der Motivation umfassen (vgl. 4.2.2 „Motivationale Faktoren“).

In den Fällen, in denen es um die Messung von Veränderungen geht, hat sich zudem die Durchführung eines Prä- sowie eines Posttests bewährt. Prätests sind beispielsweise der Vorwissenstest und die Befragungen zu den Lernvoraussetzungen, die vor dem Einsatz des Lernangebots durchgeführt werden. Die Lernerfolgskontrolle gehört zu den Posttests.

4.3.3 Lerntransfer

Unter *Lerntransfer* versteht man im weitesten Sinne erstens die Übertragung von Gelerntem aus einem Lernfeld auf ein Funktionsfeld und zweitens jede Beeinflussung einer nachfolgenden Aktivität durch vorangehendes Lernen. Berufliche Aufgaben werden zunehmend komplexer und unterliegen dynamischen Veränderungen auf Grund von neuen Technologien, neuen Märkten, neuen Angeboten etc. Es macht daher keinen Sinn, in Ausbildungsgängen die konkreten Arbeitsgänge bereits auf Vorrat im Detail zu trainieren. Vielmehr bleibt nur die Möglichkeit, grundlegende Zusammenhänge, Methoden und Verfahren mit großem Anwendungsbereich und von dauerhafter Relevanz zu lehren und die Lernenden zu befähigen, damit situationsangemessen umzugehen.

Lerntransfer lässt sich nicht durch eine einfache Lernerfolgskontrolle am Ende einer Bildungsmaßnahme messen. Er zeigt sich vielmehr erst in der konkreten Anwendungssituation am Arbeitsplatz oder im Alltag.

Lerntransfer erfordert ein hohes Maß an kognitivem Einsatz. Die Fähigkeit zur Anwendung des Gelernten garantiert noch nicht die tatsächliche Umsetzung in der Praxis. Vielfach gibt es Defizite im Hinblick auf die emotionale Bereitschaft bzw. dem Willen, das erworbene Wissen umzusetzen. Dieses Phänomen bezieht sich auf die *affektive Dimension* des Transferproblems. Danach sind häufig tiefliegende emotionale Faktoren der Grund für den fehlenden Lerntransfer. Dazu gehört z. B. die emotionale Bindung an bestimmte Software, mit der man gewohnheitsmäßig und häufig seit längerer Zeit arbeitet.

Lerntransfer hat zudem eine soziale bzw. organisationale Dimension. Die Anwendung des erworbenen Wissens verändert die Arbeitsabläufe und Organisationsformen der Arbeit und beeinflusst auch das Verhalten anderer

Mitarbeiter, die möglicherweise andere Interessen haben und den Veränderungen kritisch oder ablehnend gegenüber stehen.

Neben solchem Verhalten sind es vor allem das Führungsverhalten, Anreizsysteme, Partizipationsmöglichkeiten und Unternehmensziele, die zu den zentralen transferrelevanten Faktoren des organisationalen Umfeldes zählen. Die wohl bedeutendsten transferfördernden bzw. -hemmenden Faktoren am Arbeitsplatz liegen in der Unternehmensorganisation.

Als Erhebungsmethode zur Evaluation von Weiterbildungsmaßnahmen bietet sich das Fragebogenverfahren an, und zwar sowohl hinsichtlich der Praktikabilität im Unternehmen als auch im Hinblick auf die Diskussion um Qualität und Validität der Erhebungsinstrumente. Ein Beispiel dafür ist der LernTransfer-Evaluationsfragebogen LTE von Lemke (1995). Mit diesem Instrument können sowohl Hinweise über eine transferorientierte Gestaltung von Seminaren als auch über Gestaltungsmöglichkeiten des betrieblichen Umfelds gewonnen werden. Befragung dazu können beispielsweise nach 3-6 Monaten am Arbeitsplatz durchgeführt werden. Die Fragen können die Rolle des Vorgesetzten, der Kollegen und des Trainers, die Lernatmosphäre, motivationale Aspekte und das Arbeitsplatzumfeld nach Rückkehr an den Arbeitsplatz umfassen.

Die folgende Tabelle zeigt die Fragenbeispiele.

Sie haben an der Schulungsmaßnahme „A“ vor „X-Zeit“ teilgenommen. Nachdem einige Zeit verstrichen ist, möchten wir von Ihnen wissen, wie Sie die Schulung aus heutiger Sicht einschätzen. Bitte geben Sie jeweils an, in welchem Ausmaß Sie den Aussagen zustimmen.

trifft zu trifft eher zu trifft eher nicht zu trifft nicht zu
1 2 3 4

	1	2	3	4
Nach der Schulung war mir klar, welche Möglichkeiten der Umsetzung ich an meinem Arbeitsplatz hatte.				
Mir fiel es schwer, nach der Schulung meinem Vorgesetzten klar zu machen, was in der Schulung abgelaufen war.				
Nach der Schulung stand ich allein mit den Problemen der Umsetzung.				
Das Klima in meiner Arbeitsgruppe war einer Übertragung meiner Lernerfolge förderlich.				

Tabelle 13: Itemsauszug zur Erfassung des Lerntransfers

Die im Abschnitt 4 vorgestellten Evaluationsbereiche können mit unterschiedlichen Evaluationsmethoden erfasst und analysiert werden. Diese Methoden werden im Folgenden vorgestellt und hinsichtlich ihrer Vor- und Nachteile beschrieben. Dabei wird sowohl auf Methoden der Datenerhebung als auch der Datenauswertung eingegangen.

5 Evaluationsmethoden

5.1 Datenerhebungsmethoden

5.1.1 Dokumentenanalyse

Die *Dokumentenanalyse* erfolgt immer dann, wenn Personen oder Gruppen, die wichtige Daten liefern könnten, nicht erreichbar sind oder die Informationsentnahme aus Dokumenten effektiver ist. Mit dieser Methode werden Daten zur Präzisierung von Inhalten und Zielen, zur Entwicklung von Design-Konzepten und zur Formulierung von Kriterien gesammelt. Außerdem können sich die *Dokumentenanalysen* zum Beispiel auch auf Satzungen oder Unternehmensleitlinien beziehen. In diesem Kontext geht es vielfach darum, Informationen zu den Rahmenbedingungen für Aus- und Weiterbildung im Unternehmen sowie zu den Zielen von Bildungsmaßnahmen zu liefern.

Die Dokumentenanalyse wird als *non-reaktives* Verfahren bezeichnet. Der Terminus „non-reaktiv“ bedeutet, dass die Datenbasis – die Dokumente – durch die Teilnahme an der Untersuchung nicht verändert werden. Das unterscheidet die Dokumentenanalyse beispielsweise von Fragebogenstudien

Die Methode zeichnet sich außerdem dadurch aus, dass sie relativ preisgünstig ist und den „Lieferanten“ der Informationen keine zusätzliche Belastung auferlegt.

Dennoch liefert die Dokumentenanalyse vielfach nur einen Teil der benötigten Informationen und erfordert häufig eine zusätzliche Interpretation. Sie dient daher vorwiegend als Vor- bzw. Ergänzungserhebung.

5.1.2 Befragung

Die *Befragung* stellt die am häufigsten angewandte Methode der Datenerhebung für die Evaluation dar. Sie wird vor allem eingesetzt, wenn Informationen über die Qualität, den Nutzen und die Wirkung aus der Sicht der Beteiligten ermittelt werden sollen. Da die Befragten direkt auf die Fragen antworten, handelt es sich hierbei um ein *reaktives* Verfahren. Die Befragung kann mündlich oder schriftlich erfolgen.

Befragungen können unterschiedliche Strukturierungsgrade aufweisen. Man unterscheidet zwischen strukturiertem, halbstrukturiertem und offenem Interview.

Für ein *strukturiertes Interview* ist ein Interviewleitfaden erforderlich, der für den Interviewer verbindlich ist. Es werden präzise Fragen formuliert, die vom Befragten kurz beantwortet werden können. Strukturierte Interviews sind für klar begrenzte Themenbereiche geeignet, über die man bereits detaillierte Vorkenntnisse besitzt.

Bei einem *halbstrukturierten Interview* stehen die zentralen Fragen zwar fest, nicht jedoch deren Reihenfolge. Zudem hat der Interviewer einen Gestaltungsspielraum. Er kann – je nach Antwort auf die zentralen Fragen – aus der Situation heraus weiterführende Zusatzfragen stellen. Charakteristisch für diese Befragungsform ist ein Interview-Leitfaden, der dem Interviewer mehr oder weniger verbindlich die Art und die Inhalte des Gesprächs vorschreibt (siehe Tabelle 14).

In einem *offenen Interview* hingegen werden dem Interviewer sowohl die Formulierung der Fragen als auch deren Reihenfolge überlassen. Das offene Interview ist vor allem für explorative Studien geeignet. Solche Studien dienen zu einer ersten Orientierung in einem Forschungsfeld.

Die folgende Tabelle zeigt für ein halbstrukturiertes Interview Auszüge aus einem Interviewleitfaden.

Leitfaden zur Erfassung der Einstellungen zum Computer /Internet
Welche Einstellungen haben Sie zum Computer (nützlich, undurchschaubar, unterhaltsam, notwendig...)?
Welche Funktion hat der Computer für Sie (Arbeits- bzw. Lernmittel, Unterhaltungsmittel...)?
Was bringen Computer- (Internet-) Kenntnisse für Ihr Privat- und Berufsleben?

Tabelle 14: Leitfadenauszug zur Erfassung der Einstellung zum Computer/Internet

Bevor die mittels Befragung gewonnenen Daten ausgewertet werden können, müssen sie zunächst angemessen aufbereitet und dokumentiert werden. Das beginnt mit der Transkription der Audioaufzeichnungen. Transkripte sind dabei nicht nur eine Umschreibung der Interviewtexte. Sie können vielmehr auch die prägnanten Merkmale des Gesprächsverlaufs (z.B. Tonhöhe, Pausen und Kommentare) enthalten, die für die nachfolgende Interpretation wichtig sind. In Tabelle 15 sind verschiedene Transkriptionszeichen zusammengestellt (vgl. Bortz & Döring, 2002 S 312).

Transkriptionszeichen	Bedeutung
Großschreibung	Betonung von Silben
*	Kurze Pause
**	Längere Pause
/	Abbruch eines Wortes oder Satzes
In Klammern und Großbuchstaben	Kommentar
<	Tonhöhe fallend
>	Tonhöhe steigend

Tabelle 15: Beispiele für Transkriptionszeichen

Die mündliche Befragung ist jedoch keineswegs unproblematisch. Sie birgt beispielsweise die Gefahr der unzulässigen Beeinflussung der Befragten durch den Interviewer in sich. Zudem sind mündliche Befragungen zeitaufwendig und lassen vielfach keinen größeren Stichprobenumfang zu. In der Praxis kommt es auch vor, dass Befragte die Beantwortung aller oder einzelner Fragen ablehnen und dadurch eine sinnvolle Datenauswertung erschwert wird. Zumindest einigen der angeführten Probleme kann durch eine sorgfältige Vorbereitung der Datenerhebung und der Probanden sowie

durch die systematische Schulung der Interviewer und durch Probeinterviews entgegengearbeitet werden.

Der Vorteil dieser Methode besteht darin, dass nachgefragt werden kann, und dass wichtige Aspekte im Gespräch elaboriert werden können. Zudem können neue Aspekte, die nicht im Interviewleitfaden enthalten sind, aber von den Befragten thematisiert werden, aufgegriffen und geklärt werden.

Die *schriftliche Befragung* kann mithilfe von Fragebögen oder Checklisten durchgeführt werden. Sie erfolgt entweder postalisch oder als Online-Befragung. Entscheidend ist hierbei, dass außer einer einleitenden Instruktion keine weiteren Erläuterungen notwendig werden, da keine Rückfragen möglich sind. Daher sollte vor der eigentlichen Befragung ein Prätest durchgeführt werden, der Defizite in den Formulierungen aufdeckt.

Bei der schriftlichen Befragung wird zwischen offenen und geschlossenen Fragen unterschieden. Für die offenen Fragen sind keine Antwortkategorien vorgegeben; die Beantwortung erfolgt frei. Demgegenüber sind bei geschlossenen Fragen entweder Antwortkategorien oder abgestufte Rating-Skalen vorgegeben.

Die folgende Tabelle zeigt Beispiele mit verschiedenen Antwortmöglichkeiten.

Wo haben Sie meistens mit dem E-Learning Angebot gelernt?

Zu Hause Am Arbeitsplatz Anderswo: _____
 () ()

Wie sehr fühlten Sie sich durch folgende technische Probleme beeinträchtigt?

sehr viel eher viel eher wenig sehr wenig
 1 2 3 4

	1	2	3	4
a) Lange Ladezeit				
b) Probleme beim Einloggen				
c) Abstürzen des E-Learning-Angebotes				
d) Tempo des Internetzugangs				

Tabelle 16: Beispiele mit unterschiedlichen Antwortmöglichkeiten für die schriftliche Befragung

Die Probleme der schriftlichen Befragung ergeben sich aus der unkontrollierten Erhebungssituation. Charakteristisch sind:

- Verständnisprobleme,
- Nicht-Eindeutigkeit der Fragen,
- Suggestivfragen.

Des Weiteren ist die Antwortrate, die sogenannte „Rücklaufquote“ der Fragebögen, vielfach geringer als bei anderen Methoden. Zudem sind die Antworten bei offenen Fragen schwer zu kategorisieren und auszuwerten.

Als Vorteile der schriftlichen Befragung gelten vor allem:

- die einfache Art und Weise, mit der umfangreiche Informationen gesammelt werden können,

- die vergleichsweise geringen Kosten,
- der geringe Zeitaufwand,
- die verglichen mit anderen Methoden größere Anonymität, die im Gegensatz z. B. zu Interviews eine offenerere und ehrlichere Beantwortung der Fragen bewirken kann,
- die leichtere Koordination des Vorgehens sowie
- eine vergleichsweise objektive Dateninterpretation, die darauf beruht, dass alle Befragten den gleichen Fragebogen erhalten.

Um eine hohe Qualität der Befragung zu gewährleisten, sollten bei der Erstellung eines Fragebogens zumindest die folgenden Grundsätze beachtet werden (Hammann & Erichson, 2000).

1. Für die *Formulierung der Fragen* ist zu beachten, dass:

- das verwendete Vokabular für die Testperson verständlich und geläufig ist,
- keine vagen oder mehrdeutigen Formulierungen benutzt werden,
- Antworttendenzen in der Form, dass der Befragte einer bestimmten Antwort kaum ausweichen kann, vermieden werden,
- die Fragen allgemein beantwortbar sind und keine spezifische Sachkenntnis erfordern.

2. Im Hinblick auf den *Umfang eines Fragebogens* lässt sich kaum eine verbindliche Aussage treffen. Im Allgemeinen wird davon ausgegangen, dass monothematische Befragungen nicht länger als eine halbe Stunde andauern sollten. Polythematische Fragebögen hingegen bedürfen naturgemäß eines größeren Zeitumfangs.

3. Die *Übersichtlichkeit eines Fragebogens* kann durch eine ansprechende grafische Gestaltung erheblich erhöht werden. Fragebögen bestehen aus verschiedenen Elementen wie den Fragen, den Antwortvorgaben bzw. dem Platz für offene Antworten sowie Informationen. Diese Elemente sollten auch optisch getrennt dargestellt werden. Dies kann zum Beispiel durch eine Umrahmung oder die Wahl verschiedener Schriftarten erfolgen.

5.1.3 Beobachtung

Durch die *Beobachtung* werden Daten über sinnlich wahrnehmbare Ereignisse und Verhaltensweisen gesammelt. Dabei ist stets zu beachten, dass Wahrnehmungen die Realität keineswegs spiegelbildlich erfassen. Es kann nie alles gleichzeitig, sondern es können stets nur zentrale Aspekte beobachtet werden können. Zudem geht es darum, die Subjektivität des Beobachters so weit wie möglich zu kontrollieren.

Beobachtungen sind insbesondere angebracht, wenn zu vermuten ist, dass die verbale oder schriftliche Selbstdarstellung des Untersuchungsteilnehmers bewusst oder unbewusst verfälscht wird (Bortz & Döring, 2002). Sie kann als Fremd- oder Selbstbeobachtung durchgeführt werden.

Bei der Selbstbeobachtung betrachtet der Beobachter sich selbst. Dabei gibt es jedoch das Problem der fehlenden Distanz und der eingeschränkten Wahrnehmung.

Bei der Fremdbeobachtung werden andere Personen zum Untersuchungsgegenstand des Beobachters. Sie kann nach verschiedenen Kriterien weiter spezifiziert werden. Es wird unterschieden:

- nach der Strukturiertheit in strukturiert versus unstrukturiert,
- nach der Teilnahme des Beobachters in teilnehmend versus nicht teilnehmend und
- nach der Informiertheit der Teilnehmer der Beobachtung in offene versus verdeckte Beobachtung.

Bei der strukturierten Fremdbeobachtung sind Gegenstand, Dauer, Zeitpunkt und Zielstellung genau festgelegt. Ein Beobachtungsprotokoll dient dazu, alle Vorgänge genau zu erfassen (Greve & Wentura, 2002). Dabei ist es hilfreich, mit Kodierungen zu arbeiten, um die beobachteten Informationen unmittelbar notieren zu können.

Im Folgenden werden Beobachtungsprotokolle skizziert. Im Beobachtungsprotokoll der Abbildung 2 werden bestimmte Verhaltensweisen durch Kreuze registriert. Im Beobachtungsprotokoll der Abbildung 3 kann das beobachtete Verhalten mit unterschiedlichen Ausprägungen aufgezeichnet werden.

Person:															
Beobachter:															
Verhalten	Zeit in Minuten															
	5 10 15 20 25 30 35 40 45 50 55 60 65															
Zeigt Solidarität	<table border="1"><tr><td></td><td>x</td><td>x</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td></tr></table>		x	x												x
	x	x												x		
Stimmt zu	<table border="1"><tr><td></td><td></td><td></td><td></td><td>x</td><td>x</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>					x	x									
				x	x											
Macht Vorschläge	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>															
Äußert Meinung	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td><td></td></tr></table>											x				
										x						

Abbildung 2: Beobachtungsprotokoll

Beobachtung zum Zeitpunkt 1														
<i>Konzentriert</i>	-3	-2	-1	0	+1	+2	+3	<i>Unkonzentriert</i>						
<i>Geschickt</i>	-3	-2	-1	0	+1	+2	+3	<i>Ungeschickt</i>						
<i>Selbständig</i>	-3	-2	-1	0	+1	+2	+3	<i>Unselbständig</i>						
Beobachtung zum Zeitpunkt 2														
<i>Konzentriert</i>	-3	-2	-1	0	+1	+2	+3	<i>Unkonzentriert</i>						
<i>Geschickt</i>	-3	-2	-1	0	+1	+2	+3	<i>Ungeschickt</i>						
<i>Selbständig</i>	-3	-2	-1	0	+1	+2	+3	<i>Unselbständig</i>						

Abbildung 3: Beobachtungskontrolle im Zeitverlauf

Bei der unstrukturierten Beobachtung werden die Beobachtungskategorien erst im Nachhinein, d. h. nach erfolgter Beobachtung festgelegt. Dieses

Vorgehen ist explorativ und kann der strukturierten Beobachtung vorausgehen, um Kategorien dafür zu gewinnen.

Charakteristisch für eine teilnehmende Beobachtung ist, dass der Beobachter selbst aktiv am Geschehen beteiligt ist. Sein Vorgehen ist daher meist unstrukturiert. Demgegenüber nimmt der Beobachter bei der nicht-teilnehmenden Beobachtung eine passive Rolle ein, wodurch ihm ermöglicht wird, eine strukturierte Beobachtung durchzuführen.

Die Fremdbeobachtung kann in offener oder in verdeckter Form durchgeführt werden. Als „offen“ wird die Form bezeichnet, wenn den untersuchten Personen die Rolle des Beobachters bekannt ist. Bei der verdeckten Form hingegen soll die Beobachtung unbemerkt bleiben. Sie findet zum Beispiel hinter sogenannten „Einwegscheiben“ statt, die nur für den Beobachter transparent sind.

Eine Beobachtung kann nicht nur visuell, das heißt mit dem menschlichen Auge sondern auch mittels instrumenteller oder psychographischer Verfahren durchgeführt werden.

Unter instrumenteller Beobachtung versteht man zum Beispiel die Erhebung von Daten mittels Videokamera, elektronischer Zähl- und Messeinrichtungen, Scanning oder Telemeter.

5.1.4 Verhaltensrecording

Das *Verhaltensrecording* wird auch als Benutzerprotokoll bezeichnet. Es wird mithilfe von Logfiles durchgeführt und dient der Erfassung von Interaktionen mit einem Softwareprogramm. Wie detailliert diese Protokollierung durchgeführt wird, ist vom Zweck abhängig, der damit verfolgt wird. Die Protokollierung kann sich vom Mouseclick über das Aufrufen von Menüpunkten bis hin zur Aufzeichnung aller durchgeführten Aktionen erstrecken (Gunn, 2002). Die nachfolgende Tabelle 16 stellt beispielhaft ein solches Benutzerprotokoll dar.

Lerner 1

Daten im April 2002

Zugriffszeitpunkt	Bearbeitetes Element	Kurzbeschreibung des Elements	Zugriffshäufigkeit	Verweildauer
10.04.2002 09:01:55	Informationsmanagement im Electronic Business	Es werden Themen zum Electronic Business behandelt. Hauptaugenmerk wird auf das Informationsmanagement gelegt.	9	0:19:05
10.04.2002 09:21:00	Studienangelegenheiten	Hier kann eine Aktivitätsbeschreibung stehen.	13	0:00:09
10.04.2002 09:21:09	Der Begriff eCommerce	Aktivitätsbeschreibung	5	0:05:11
10.04.2002 09:01:55	Komplexe Problemstellung	Aktivitätsbeschreibung	7	0:15:26

Tabelle 17: Benutzerprotokoll

Das Verhaltensrecording ist eine objektive Methode, die relativ zuverlässige statistische Daten liefert. Es ist jedoch nur dann sinnvoll, wenn Verhaltensdaten auf einem sehr feinkörnigen Niveau erfasst werden sollen und wenn unter Verwendung anderer Methoden zusätzliche Daten herangezogen werden können. Aus der bloßen Protokollierung wird nicht ersichtlich, warum sich die jeweiligen Benutzer entsprechend verhalten haben.

5.1.5 Test

Tests dienen zur Messung abgrenzbarer Verhaltens- und Leistungsmerkmale. Das Ziel dieses Verfahrens ist es, den Ausprägungsgrad eines Merkmals zu erfassen. Dies kann einerseits in Form einer Querschnittsdiagnose geschehen. Dabei wird die Stellung des Einzelindividuums innerhalb einer vergleichbaren Gruppe hinsichtlich bestimmter Leistungs- oder Persönlichkeitsmerkmale getestet. Andererseits kann eine Längsschnittdiagnose erfolgen. Hier werden die Merkmalsänderungen eines Individuums oder mehrerer Gruppen innerhalb definierter Zeitspannen verfolgt. Anschließend erstellt man ein Verlaufsprofil (Lienert & Raatz, 1994).

Tests laufen in der Regel standardisiert ab. Störende Einflüsse können dabei bewusst kontrolliert und vermieden werden.

Tests müssen 3 Hauptgütekriterien erfüllen:

- Objektivität,
- Reliabilität (Zuverlässigkeit) und
- Validität (Gültigkeit).

Die Objektivität besagt, dass verschiedene Evaluatoren zu den gleichen Interpretationsergebnissen gelangen. Die Reliabilität ist gegeben, wenn bei einer Wiederholung des Tests unter den gleichen Bedingungen vergleichbare Ergebnisse erzielt werden. Als „valide“ bezeichnet man einen Test, wenn er auch misst, was er zu messen vorgibt. Zudem sollte ein Test den Nebengütekriterien der Normierung, Vergleichbarkeit, Ökonomisierung und Nützlichkeit genügen.

Für Tests bestehen sehr vielfältige Klassifizierungsmöglichkeiten. Sie sind beispielsweise nach der Art der Bezugsgröße in normorientierte und kriteriumsorientierte Tests zu unterscheiden. Bei den normorientierten Tests wird die individuelle Testleistung mit der einer Vergleichsgruppe verglichen. Bei kriteriumsorientierten Tests hingegen wird die individuelle Testleistung durch Vergleich mit vorher festgelegten Kriterien gemessen.

Tests können nach der Art des zu erfassenden Merkmals in Intelligenz-, Leistungs- und Persönlichkeitstests unterschieden werden.

Testaufgaben bestehen grundsätzlich aus zwei Elementen, dem Problem bzw. der Frage und der Problemlösung bzw. der Schlüsselantwort.

Zusätzlich gibt es situative Tests. Beispiele dafür sind Arbeitsproben, Rollenspiele, Fallstudien, Gruppendiskussionen, Planspiele oder Simulationen. Sie dienen der Erfassung von Vorgehensweisen zur Bewältigung realitätsnaher Aufgaben.

5.1.6 Expertenbeurteilung

Bei einer *Expertenbeurteilung* bewertet ein Experte die Qualität eines fertig gestellten Produkts, zum Beispiel einer Lernsoftware. Dabei bedient er sich verschiedener Checklisten oder Kriterienkataloge. Unter Checklisten versteht man ad hoc zusammengestellte, kurze Listen. Demgegenüber werden Kriterienkataloge vielfach für einen bestimmten Zweck systematisch konzipiert. Im Kontext computer- und netzbasierter Lernangebote sollen solche Kataloge insbesondere die medienpädagogische, technische und ökonomische Qualität des Produkts beurteilen helfen. Kriterienkataloge dienen sowohl der Bewertung einzelner Kategorien als auch der Analyse der Gesamtqualität eines Produkts. Dabei ist auf die richtige Gewichtung der einzelnen Kriterien und der Kriteriumskategorien zu achten. Hierbei wird zwischen primären und sekundären Kriterien unterschieden. Diese werden individuell je nach Untersuchungsziel festgelegt und gehen entsprechend differenziert in die Bewertung ein (Fricke, 2000).

Auch die Anwendung dieser Kriterienkataloge muss den Gütemaßstäben der Objektivität, Reliabilität und Validität gerecht werden, um eine entsprechende Qualität der Expertenbeurteilung zu gewährleisten. Beispielsweise wird die Reliabilität durch charakteristische Beurteilungstendenzen (z.B. Instabilität im Zeitverlauf) gemindert, die auch für Experten relevant sind: Die Objektivität ist beispielsweise eingeschränkt, wenn verschiedene Experten nicht in genügendem Maße in ihrer Beurteilung übereinstimmen.

Die mangelnde Beurteilerübereinstimmung ergibt sich vor allem daraus, dass die Experten von verschiedenen Annahmen ausgehen und unterschiedliches Hintergrundwissen besitzen. Aus diesem Grund ist ein gezieltes Beurteilertraining unerlässlich. Zudem hat sich als günstig erwiesen,

die Expertenbeurteilung mittels Kriterienkatalogen mit empirischen Befunden zu kombinieren.

Die Bewertungsfragen können unterschiedlich skaliert werden. Beispielsweise sind bei Zwei-Punkt-Skalen die Antworten „ja“ oder „nein“ möglich. Die Items können aber auch je nach dem Grad des Zutreffens mit mehr als drei Stufen skaliert sein.

Bei der Arbeit mit Kriterienkatalogen sollten bereits bestehende Instrumente nicht einfach übernommen, sondern in einem aktiven Prozess auf die eigenen Erfordernisse und Zwecke hin überprüft und gegebenenfalls angepasst werden.

Beispiele für Kriterienkataloge

Der Kriterienkatalog MEDA (Gräber, 1990)

Ein Beurteilungsinstrument für pädagogische Lernsoftware im Bereich der Erwachsenenbildung ist zum Beispiel MEDA. Das Ziel des Kriterienkatalogs ist es, einen methodischen Rahmen für die Evaluation vorzugeben, der entsprechend den unterschiedlichen Vorstellungen der Beurteiler individuell angepasst werden kann (Meier, 2000). Dazu dient eine hierarchische Struktur, die durch die nachfolgende Abbildung verdeutlicht wird.

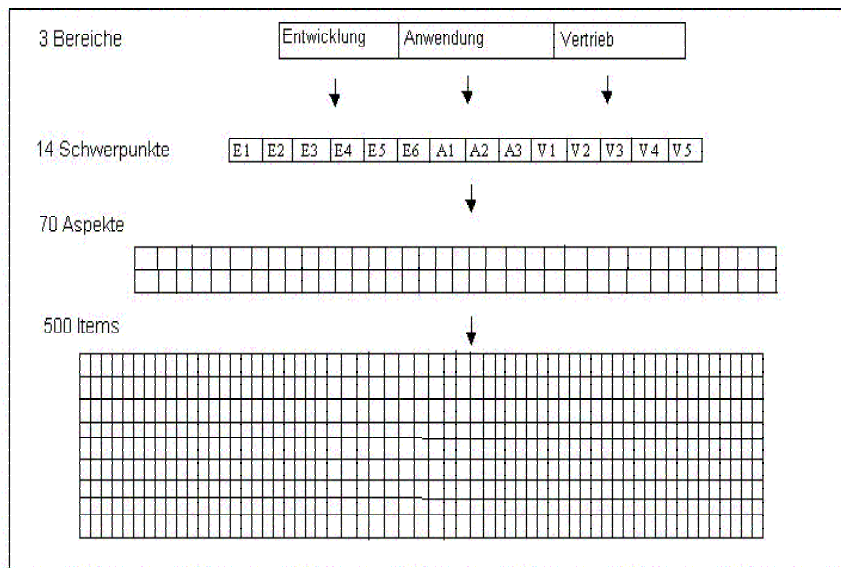


Abbildung 4: Hierarchische Struktur vom Kriterienkatalog MEDA

Der Kriterienkatalog besteht aus insgesamt 500 Fragen zu computer- und netzbasierten Lernangeboten, die mit „Ja“ oder „Nein“ zu beantworten sind. Der Evaluator kann je nach Bewertungsziel und Intention die für ihn relevanten Fragen auswählen. Unterstützt wird er dabei durch einen dreistufigen Filter bestehend aus „Bereiche“, „Schwerpunkte“ und „Aspekte“.

Der Kriterienkatalog AKAB

AKAB ist ein elektronischer Kriterienkatalog, der für die Beurteilung von Lernsoftware, insbesondere von Computer-Based-Trainings, in der Automobilindustrie entwickelt wurde. Der Katalog besteht aus 74 Kriterien, die wiederum den folgenden 13 Kriteriengruppen zugeordnet sind:

Kurzbeurteilung – Fachliche Beurteilung – Interaktion – Rückmeldung – Hilfen – Lernzielkontrollen – Motivation – Ablaufsteuerung – Dokumentation – Textgestaltung – Bildschirmgestaltung – Multimediaeinsatz – Inhalt.

Alle Kriterien werden innerhalb einer vierstufigen Ratingskala mit ja /eher ja /eher nein /nein bewertet. Jede Frage wird zusätzlich einer von den Autoren festgelegten und nicht veränderbaren Gewichtung mittels Faktoren von 1 bis 4 unterzogen. Dies erlaubt eine weitere Programmdifferenzierung. Anschließend wird für jede Frage das Produkt aus der Benotung und der entsprechenden Gewichtung gebildet (Meier, 2000). Ein Beispiel dafür ist in der folgenden Abbildung dargestellt.

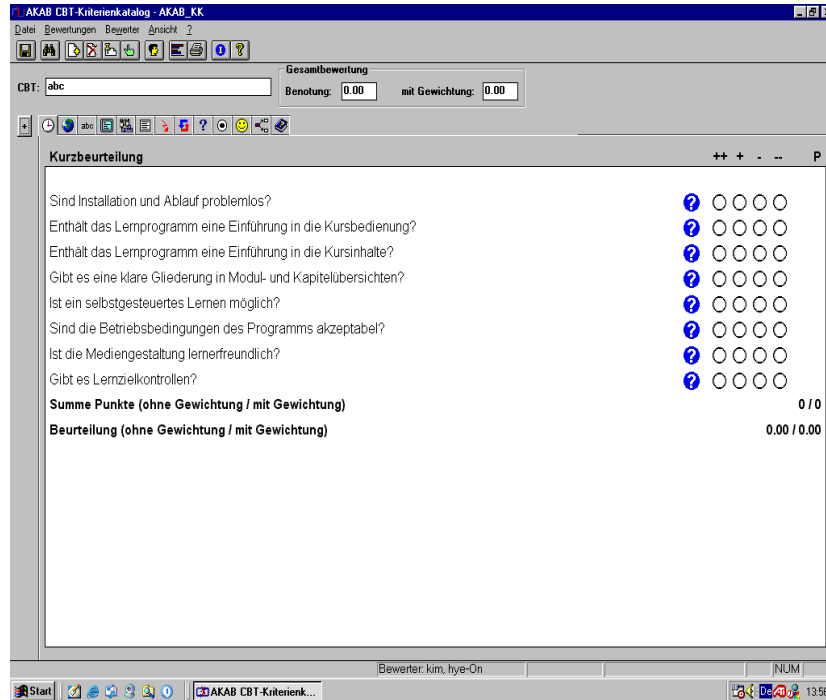


Abbildung 5: Bewertungsbildschirm von AKAB

AKAB ist durch die elektronische Erfassung und Auswertung sowie durch den eher geringen Umfang für den Benutzer überschaubar und relativ unkompliziert anzuwenden. Dem stehen jedoch deutliche inhaltliche Mängel gegenüber. Das Instrument kann der Struktur und dem Charakter der jeweiligen Lernsoftware nicht angepasst werden, d.h., es können keine weiteren Kriterien aufgenommen oder weniger und nicht relevante Kriterien

gelöscht werden. Damit ist AKAB lediglich für jene Selbstlernprogramme anwendbar, für die es konstruiert worden ist.

5.2 Datenauswertung

Nachdem die Daten mit unterschiedlichen Methoden erhoben worden sind, werden sie ausgewertet. Idealerweise wird die Art und Weise der Datenanalyse vor der Datenerhebung festgelegt, nur so ist eine optimale Kompatibilität zwischen dem Ziel der Untersuchung, der Datengewinnung und der Datenauswertung gewährleistet. Für diesen Prozess der Abstimmung stehen verschiedene Modelle zur Verfügung. (Wottawa, 2001).

Das *allgemeinpsychologische Modell* geht davon aus, dass alle Personen nach dem gleichen Funktionsmodell beschreibbar sind. Die Ergebnisse der Einzelpersonen unterscheiden sich nur durch Zufallsschwankungen. Dieses Modell wird vor allem verwendet, wenn der mittlere Effekt der einzelnen Maßnahmen interessiert.

Das *differentialpsychologische Modell* geht davon aus, dass zwar für alle Personen im Prinzip das gleiche Funktionsmodell vorliegt, die Personen sich aber systematisch hinsichtlich der Modellparameter unterscheiden können. So wird zum Beispiel bei der Durchführung eines konkreten E-Learning-Kurses davon ausgegangen, dass die Teilnehmer unterschiedliche Ausgangsniveaus besitzen.

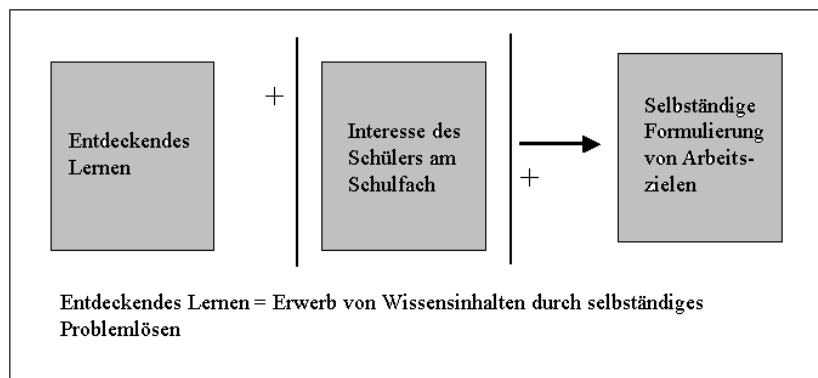
Das *individualpsychologische Modell* geht davon aus, dass für jede Person unterschiedliche Variablen relevant und in unterschiedlicher Weise miteinander verknüpft sind. Evaluationen beziehen sich auf individuumsspezifische Messungen. Da solche Datenerhebungen sehr aufwendig sind, ist der Einsatz dieses Modells nur selten möglich.

Die mit verschiedenen Methoden erhobenen Daten werden qualitativ oder quantitativ ausgewertet. Mit qualitativen Auswertungsverfahren wird verbales und nichtnumerisches Material bearbeitet. Demgegenüber werden die numerischen Daten mit statistischen Verfahren ausgewertet. Interviewtranskripte oder Beobachtungsprotokolle können zum Beispiel sowohl quantitativ als auch qualitativ analysiert werden.

5.2.1 Qualitative Auswertungsverfahren

Ziel der qualitativen Inhaltsanalyse ist es, die manifesten und latenten Inhalte in ihrem Bedeutungsfeld zu interpretieren. Die Inhaltsanalyse geht dabei in intersubjektiv nachvollziehbaren Arbeitsschritten vor. Gültige Interpretationen müssen konsensfähig sein. Es gibt zahlreiche Varianten qualitativer Inhaltsanalyse. Ein Beispiel dafür sind die Strukturierungshilfen.

Anhand von *Strukturierungshilfen* (zum Beispiel Struktur-lege-Technik, Scheele & Groeben, 1988) kann komplexes Material ausgewertet bzw. strukturiert werden. Dazu gehören zahlreiche Verfahren der Inhaltsanalyse, die insbesondere bei der Auswertung offener Interviews hilfreich sein können. Zur Auswertung der Interviewinhalte werden häufig Kärtchen mit den wichtigsten inhaltlichen Konzepten verwendet. Die Konzeptkärtchen werden mit Hilfe von Pfeilkärtchen, die Ursache-Wirkungs-Zusammenhänge darstellen, zu einer pfadanalytischen Struktur verbunden. Dabei können beispielsweise folgende Relationen angewendet werden: „definitorisch gleich“, „Über-/Unterbegriff(e)“, „Und-Verbindungen“ und „Oder-Verbindungen“. Die so gewonnene subjektive Struktur kann insbesondere bei relativ unerforschten Problembereichen zu einer Erweiterung, Differenzierung oder Vereinfachung von Annahmen führen. Die folgende Abbildung zeigt eine Anwendung der Struktur-lege-Technik.



(Quelle: Scheele & Groeben, 1988, S.59)³

³ Die Abbildung zeigt, dass entdeckendes Lernen zur Fähigkeit, selbständig Aufgabenziele zu formulieren, führt, jedoch nur dann, wenn der Schüler

Abbildung 6: Anwendungsbeispiel der Struktur-Lege-Technik

An die qualitative Auswertung kann sich auch eine quantitative Bearbeitung des Datenmaterials anschließen. Insgesamt sollten so gewonnene Erkenntnisse jedoch immer nur formativ und nicht als summatives Ergebnis dargestellt werden.

5.2.2 Quantitative Auswertung

Vor Beginn einer quantitativen empirischen Untersuchung ist zu klären, auf welchem Messniveau die Daten erhoben werden sollen. Daraus ergeben sich vielfach weitreichende Konsequenzen für die Auswertung. Es empfiehlt sich eine Gliederung nach dem Skalenniveau. Hierbei werden unter anderem zwei Skalentypen unterschieden (Bortz & Döring, 2002): Nominal- und Intervallskalen.

Skalenniveaus

Bei Nominalskalen werden Zahlen zu diskontinuierlichen Folgen von Tatbeständen zugeordnet. Dabei geht es um Äquivalenzen. Das heißt, Objekte mit identischen Merkmalsausprägungen erhalten identische Zahlen. Um welche Zahlen es sich dabei handelt, ist beliebig. Sie beinhalten keine quantitative Aussage sondern dienen lediglich der Bezeichnung sich gegenseitig ausschließender Kategorien.

Soll zum Beispiel in einem Fragebogen das Geschlecht der Probanden erfragt werden, bietet sich dafür folgende typische Frage-Antwort-Möglichkeit auf Nominalskalenniveau an:

- a) weiblich (1)
- b) männlich (2)

an dem vorliegenden Schulfach Interesse hat. Entdeckendes Lernen ist definitorisch identisch mit „Erwerb von Wissensinhalten durch selbständiges Problemlösen“.

Statistische Operationen bei nominalskalierten Merkmalen beschränken sich in der Regel auf so genannte Häufigkeitsverteilungen, auf die im Verlauf der Lektion noch näher eingegangen wird.

Intervallskalen messen auf metrischem Niveau. Eine Intervallskala ordnet Objekten Zahlen zu, die so geartet sind, dass die Abstände zwischen den Zahlenwerten den Abständen der Stärke einer Eigenschaft entsprechen. Mit Intervallskalen können sinnvoll Differenzen, Summen oder Mittelwerte berechnet werden (vgl. nachstehendes Beispiel).

Wie sehr fühlten Sie sich durch folgende technische Probleme beeinträchtigt?						
sehr viel	eher viel	eher wenig	sehr wenig			
1	2	3	4			
			1	2	3	4
a) Lange Ladezeit						
b) Probleme beim Einloggen						
.....						

(Quelle Kim & Klauser, 2004)

Tabelle 18: Itembeispiel für die Intervallskala

Die korrekte Anwendung statistischer Verfahren zur Auswertung der gesammelten Daten setzt zunächst voraus, dass die Funktion der Statistik für die Evaluationsarbeiten richtig eingeschätzt wird. So können zum Beispiel die deskriptive Statistik und die Inferenzstatistik differenziert angewendet werden.

Die deskriptive Statistik beschränkt sich weitgehend auf die Ordnung und Zusammenfassung von Daten. Ihr Hauptziel ist die übersichtliche und anschauliche Informationsaufbereitung. Im Unterschied dazu ermöglicht die

Inferenzstatistik die Überprüfung von Hypothesen. Mit ihrer Hilfe werden theoretische Aussagen angesichts der beobachteten Realität getestet.

Deskriptive Statistik

Vor Beginn der Auswertungsarbeiten muss die testtheoretische Brauchbarkeit der Daten überprüft werden. Hierbei geht es insbesondere um die Objektivität und Reliabilität der empirischen Basis. Genügen die Daten den Anforderungen, werden sie in übersichtlicher Form als Graphiken, Tabellen oder einzelne Kennwerte zusammengestellt.

Zur tabellarischen Aufbereitung werden die Messwerte zunächst zu Kategorien zusammengefasst. Durch Auszählung werden Häufigkeiten für die einzelnen Kategorien ermittelt. Sollen zwei unterschiedlich große Kollektive hinsichtlich ihrer Merkmalsverteilung verglichen werden, können die absoluten Häufigkeiten der Kategorien auch als Prozentwerte ausgedrückt werden. Die kumulierten Prozente müssen sich zu 100% addieren.

Folgendes Beispiel zeigt eine Häufigkeitstabelle für die Nutzung des Textverarbeitungsprogramms von Studenten wirtschaftlicher Fachrichtungen.

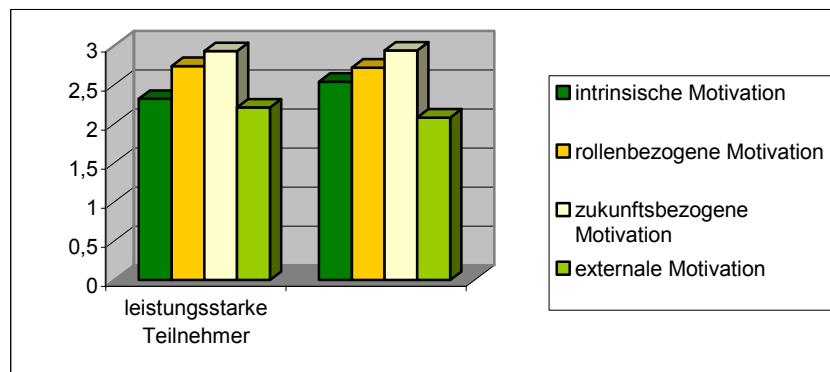
	Häufigkeit	Prozente	Kumulierte Prozente
sehr häufig	168	30,5	30,5
häufig	238	43,3	73,8
selten	108	19,7	93,5
gar nicht	36	6,5	100,0
Gesamt	550	100,0	

(Quelle: Klauser & Kim, 2003b)

Tabelle 19: Tabellarische Darstellung der Häufigkeitsverteilung

Der arithmetische Mittelwert – kurz: der Mittelwert – ist das gebräuchlichste Maß zur Kennzeichnung der zentralen Tendenz einer Verteilung. Er wird per definitionem von allen Messwerten beeinflusst. In der

folgenden Abbildung sind die Mittelwerte unterschiedlicher Motivationen von Untersuchungsteilnehmern graphisch dargestellt.



(Quelle: Klauser, Kim & Pollmer, 2003)

Abbildung 7: Graphische Darstellung der Mittelwerte

Die einzelnen Werte weichen zumeist mehr oder weniger deutlich von dem errechneten Mittelwert ab. Wird der kleinste auftretende Wert vom größten subtrahiert, so ergibt sich der absolute Streubereich. Je kleiner dieser ist, desto homogener ist die Stichprobe. Ein gebräuchliches Maß zur Kennzeichnung der Variabilität einer Verteilung ist jedoch nicht der Streubereich, sondern die *Standardabweichung*, die aus den Abweichungen der einzelnen Werte vom Mittelwert bestimmt wird. Bei der Ermittlung der Standardabweichung werden alle Werte (und nicht nur, wie beim Streubereich, die Extreme) berücksichtigt. Dieses Vorgehen ermöglicht eine besonders informationsreiche Beschreibung der gesamten Variabilität aller Werte.

Der folgende Auszug aus einer tabellarischen Darstellung zeigt in exemplarischer Weise Mittelwerte und Standardabweichungen bei der Erfassung der Wichtigkeit der Kriterien für ein gutes E-Learning-Angebot.

	Mittelwert	Standard- abweichung
Übersichtlichkeit	2.95	.77
Klare Lernzielformulierung	2.72	.79
Wahlmöglichkeit der Lerneinheit	3.09	.77
Möglichkeit, das Lerntempo selbst zu bestimmen	2.92	.84
.....		

(Quelle Klausner & Kim 2003b)

Tabelle 20: Auszug aus einer tabellarischen Darstellung von Mittelwerten und Standardabweichungen

Inferenzstatistik

Ausgangspunkt der Inferenzstatistik ist die Formulierung einer allgemein gültigen Hypothese. Anhand von Beobachtungen und Daten (d.h. „aufbereiteten Beobachtungen“) wird dann die „empirische Bewährung“ dieser Hypothese überprüft. Hypothesenprüfende Untersuchungen können mit vielfältigen Methoden der Inferenzstatistik durchgeführt werden:

Chi-Quadrat-Tests (bzw. χ^2 -Methoden) dienen der Prüfung von Hypothesen über Häufigkeitsverteilungen. Alle Chi-Quadrat-Verfahren laufen auf einen Vergleich von beobachteten und erwarteten Häufigkeiten und die Berechnung eines Chi-Quadrat-Koeffizienten hinaus. Aus der tabellarischen Darstellung der χ^2 -Verteilung kann nun abgelesen werden, wie wahrscheinlich ein zufälliges Zustandekommen dieses Wertes ist. Die Wahrscheinlichkeit wird durch einen p-Wert angegeben (s. folgende Tabelle). Je kleiner der p-Wert ist, desto geringer ist die Wahrscheinlichkeit, dass der Unterschied zwischen beobachteten und erwarteten Häufigkeiten zufallsbedingt ist. Für die Wahrscheinlichkeit p sind in den

Sozialwissenschaften per Konvention die Werte .05, .01 und .001 (häufig auch als 5%, 1% bzw. 0,1% bezeichnet) festgelegt.

Tabelle 21 zeigt als Beispiel die mittels einer Chi-Quadrat-Berechnung gewonnenen Ergebnisse zum Einfluss von Semesteranzahl und Geschlecht auf verschiedene Vorerfahrungen mit dem E-Learning im Rahmen des IMPULS-Projekts.

	Anzahl	%	χ^2	
			Semester	Geschlecht
Erfahrung mit E-Learning	317	60,2	3,13*	9,68**
CD-ROM	271	50,8	4,01*	8,00**
Netzbasieretes Lernen	134	25,1	0,01	17,13***
.....				

*p<.05 **p<.01 ***p<.001 (Quelle Klauser & Kim 2003b)

Tabelle 21: Auszug aus einer tabellarischen Darstellung einer χ^2 -Analyse

In empirischen Untersuchungen interessiert oftmals der Einfluss verschiedener Merkmale (unabhängiger Variablen) auf eine intervallskalierte abhängige Variable. Für solche Auswertungen kann die Varianzanalyse eingesetzt werden. Je nach der Anzahl der unabhängigen Variablen sind verschiedene Versionen der Varianzanalyse zu unterscheiden. Die einfaktorielle Varianzanalyse überprüft, dem Namen entsprechend, die Auswirkung *einer* unabhängigen Variablen auf die abhängige Variable. Wenn aufgrund inhaltlich-theoretischer Erwägungen angenommen wird, dass die Wirkungsweise *mehrerer* unabhängiger Variablen die abhängige Variable beeinflussen kann, wird die mehrfaktorielle Varianzanalyse durchgeführt. Die mehrfaktorielle Varianzanalyse bietet auch die Möglichkeit, Wechselwirkungen (Interaktionen) zwischen den unabhängigen Variablen zu überprüfen.

Tabelle 22 zeigt Ergebnisse zum Einfluss der Semesteranzahl und des Geschlechts auf die Computer- und Internetnutzung, ausgewertet mit der zweifaktoriellen Varianzanalyse. „M“ steht in der Tabelle für den Mittelwert

und „SD“ für die Standardabweichung. Zugleich wird überprüft, ob zwischen den beiden Einflussfaktoren Semesteranzahl (A) und Geschlecht (B) signifikante Wechselwirkungen auftreten (in der Tabelle als A*B bezeichnet).

	M	SD	F	A*B
			Semester (A)	Geschlecht (B)
Textverarbeitung	2,98	0,87	30,27***	1,03
Datenbankprogramm	1,58	0,78	0,03	18,91***
Tabellenkalkulation	1,80	0,84	64,01***	0,19
Präsentationsprogramm	3,20	0,85	65,03***	0,32
E-mail	3,69	0,55	2,87	0,87
Informationsrecherche	3,34	0,80	12,29***	0,02
Multimedia-Anwendung	2,15	1,03	1,79	39,57***
.....				8,23**

*p<.05 **p<.01 ***p<.001 (Quelle: Klauser & Kim, 2003b)

Tabelle 22: Auszug aus einer tabellarischen Darstellung der Ergebnisse einer zweifaktoriellen Varianzanalyse

Die Interaktion zwischen Semesteranzahl und Geschlecht in der Tabelle sagt aus, dass der Geschlechterunterschied für die Multimedia-Anwendungen bei Studierenden im Grundstudium noch deutlicher als bei Studierenden im Hauptstudium ausfällt. Abbildung 7 veranschaulicht diesen Sachverhalt.

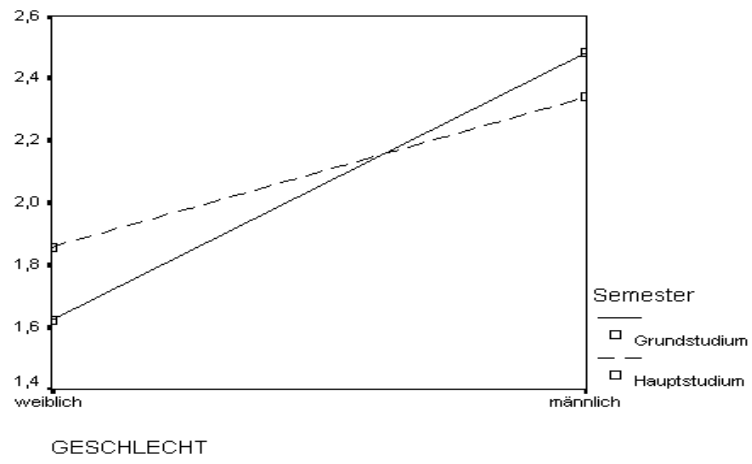


Abbildung 8: Graphische Darstellung der Interaktion einer zweifaktoriellen Varianzanalyse

6 Berichterlegung

Der letzte Schritt des Evaluationsvorhabens ist die Berichterlegung. Während Datenerhebung und -auswertung von einem wissenschaftlich vorgebildeten Evaluator ausgeführt werden sollten, liegt die Umsetzung der Ergebnisse in praktisches Handeln bei den Entscheidungsträgern bzw. Auftraggebern. Eine unmittelbare Umsetzung ist jedoch nur dann möglich, wenn eine eindeutige Verkettung zwischen empirischen Ergebnissen und daraus abzuleitenden Entscheidungen erreicht werden kann. Da dies nur selten der Fall ist, spielt die Berichterlegung des Evaluators für die Personen, die letztendlich eine Entscheidung treffen, eine wichtige Rolle.

Die Ergebnisdarstellung kann sowohl schriftlich als auch mündlich erfolgen. Mögliche Präsentationsformen von Evaluationsergebnissen sind der Fachbericht, kommentierte Zusammenfassungen, Workshops mit Mitarbeitern im Unternehmen, Poster, Broschüren und Diskussionen. Zunächst sollte sich der Evaluator also darüber im Klaren sein, welche Art von Berichterstattung von ihm gefordert wird. Der Bericht birgt die Chance, die Verantwortlichen über das Projekt zu informieren und zu verdeutlichen, warum es erfolgreich war bzw. Gründe zu nennen, die den Erfolg verhindert haben.

Der Evaluator muss genau überlegen, welche Teile der umfangreichen Ergebnisse, die bei der Evaluation gewonnen worden sind, in den Bericht aufgenommen werden sollen. Jeder Berichterlegung geht daher eine sorgfältige *Adressatenanalyse* voraus. Dabei sollten die unterschiedlichen Interessen der Adressaten beachtet werden (Mogey, 2003).

In diesen Zusammenhang ist zu entscheiden, ob sich die Berichte an Laien oder Wissenschaftler richten. So sind für Laien umfangreiche Darstellungen meist zwecklos, da höchstens Kurzzusammenfassungen, die auf der Grundlage der empirischen Fakten nachvollziehbar sind, wirklich Beachtung finden. Bei formativen Evaluationen in praxisorientierten Projekten ist ohnehin die enge Kooperation zwischen dem Evaluator und den Beteiligten gewährleistet.

Die Erarbeitung von Publikationen erfolgt in der Regel nach vorgegebenen Richtlinien. Ein Evaluationsbericht sollte einen bestimmten Informations-

umfang nicht überschreiten. Die Berichterstattung sollte in jedem Fall aktuell sein und rechtzeitig erfolgen und in einer klaren Sprache und mit übersichtlichen Visualisierungen vorgelegt werden. Es sollten sowohl qualitative als auch quantitative Daten Verwendung finden. Neben der klassischen Publikation kann die schriftliche Berichtlegung in Form eines Artikels in einem Journal, eines Memo, eines Briefs oder als Webpage erfolgen.

Insgesamt ist die Veröffentlichung von Informationen und Ergebnissen ein Kommunikationsmedium, durch das die Evaluation erst einen wesentlichen Teil ihrer Wirkung erhält. Damit der Bericht überhaupt gelesen werden kann, müssen die Ergebnisse zusätzlich in eine rezipierbare, kurze Form gebracht werden. Dafür ist meist eine separate Kurzdarstellung und Beschlussvorlage hilfreich.

In den vorangegangenen Abschnitten wurde gezeigt, wie die Daten erhoben, ausgewertet und analysiert und die Ergebnisse im Bericht dargelegt werden können. Weitgehend unberücksichtigt sind die zahlreichen Probleme geblieben, die bei der Planung und Durchführung von Evaluationsvorhaben auftreten können.

7 Probleme bei der Evaluation

Häufig beobachtete Probleme bei Evaluationsvorhaben sind zum Beispiel die mangelnde Akzeptanz der Evaluation, die Heterogenität der Zielvorstellungen und die fehlende Bereitschaft, die Ergebnisse der Evaluation anzunehmen.

Mangelnde Akzeptanz der Evaluation

Für die mangelnde Bereitschaft zur Durchführung der Evaluation kann es unterschiedliche Gründe geben. Tabelle 23 zeigt einige Auszüge möglicher Einwände gegen eine Evaluation aus verschiedenen Perspektiven im Unternehmen:

Geschäftsleitung	Es entstehen Kosten. Beteiligte Mitarbeiter könnten sinnvoller eingesetzt werden. Es könnten Unruhen im Unternehmen auftreten.
Personalabteilung	Es entsteht Mehrarbeit. Es könnten Tabubereiche angesprochen werden.
Mitarbeiter	Es besteht Unsicherheit darüber, was mit den persönlichen Daten geschieht.

Tabelle 23: Gründe für mangelnde Akzeptanz der Evaluation (Auszug)

Daher ist es wichtig, bei der Planung die unterschiedlichen Meinungen zu integrieren und sowohl Ziele als auch Vorgehensweise der Evaluation offen zu legen.

Heterogenität der Zielvorstellungen

Bei der Planung der Evaluation geraten oft die Interessen der Beteiligten (zum Beispiel die persönlichen Interessen der Schulungsteilnehmer, die organisatorischen Interessen der Personalabteilung und die wirtschaftlichen Interessen der Geschäftsführung) in Konflikt miteinander. Zur Optimierung der Untersuchung ist es notwendig, die Interessen aller Gruppen so weit wie möglich zu berücksichtigen.

Daher sind genaue Zielbeschreibungen aus der Sicht der Beteiligten und das Finden eines Konsenses vordringlich. Zudem sind die wichtigsten Evaluationskriterien und deren Grenzwerte in objektiver Form festzulegen, um die Ergebnisse in einem Ist-Soll-Vergleich auswerten zu können. Die Aufgabe des Evaluators ist es, die Beteiligten insbesondere bei der Klärung des Evaluationsziels zu unterstützen.

Tabelle 24 kann als Grundlage zur Klärung der Zielvorstellungen und zur Konsensfindung im Unternehmen dienen.

	Geschäfts- -führung	Personal- -abteilung	Mit- -arbeiter
Organisatorische Interessen				
Evaluationsziele und Kriterien				
Methodenvorzüge				
Grad des Engagements				
....				

Tabelle 24: Grundlage zur Klärung der Zielvorstellungen

Fehlende Bereitschaft zur Akzeptanz der Evaluationsergebnisse

Bei potentiell negativen Auswirkungen für Entscheidungsträger können die Evaluationsergebnisse in verfälschender Weise zusammengestellt und „unerwünschte“ Ergebnisse als „wenig aussagekräftig“ abgewertet werden. In beiden Fällen liegt eine mangelnde Bereitschaft zur Akzeptanz der Evaluationsergebnisse vor.

Zudem ist es bei größeren Projekten unmöglich, alle durch die Datenauswertung erzielbaren Ergebnisse wirklich zu verarbeiten. Daher besteht bei der Berichterlegung am Ende die Gefahr der absichtlichen oder auch unbeabsichtigten Manipulation durch die Datenauswahl. Aus diesem

Grunde ist es sinnvoll, die erhobenen Daten für Sekundäranalysen durch „kritische“ Kollegen bereitzustellen (Schweer, 1998).

8 Schlussbemerkungen

Die Planung der Evaluation ist ein komplexer Prozess. Für umfangreiche Evaluationsprojekte empfiehlt es sich, die notwendigen Teilschritte in Form eines *vorläufigen Arbeitsplanes* zu bestimmen. Dazu gehören die einzelnen Arbeitsschritte in ihrer zeitlichen Abfolge mit Angabe des erwarteten Aufwands sowie den Entscheidungen, die im Arbeitsverlauf mit dem Auftraggeber getroffen werden müssen. Hinzu kommen möglichst exakte Termin- und Kostenangaben.

Evaluationsvorhaben sind trotz sorgfältiger Planung vielfach mit Problemen verbunden. Deshalb werden insbesondere größere Projekte mit einer regelmäßigen Rückmeldung über Zeitplanung, Kostenrahmen und vereinbarte Qualitätsstandards durchgeführt. Typische *Störfälle* sind zum Beispiel Personalwechsel beim Evaluationsauftraggeber, der zu unterschiedlicher Schwerpunktsetzung führt; oder fehlende Bereitschaft der Probanden zur Mitarbeit, etwa wenn diese sich weigern, als Interviewpartner zur Verfügung zu stehen. Um rechtzeitig auf Störfälle reagieren zu können, muss der Evaluationsverlauf ständig kontrolliert werden.

Ergänzend soll an dieser Stelle eine Schwierigkeit des Untersuchungsdesigns von Evaluationsprojekten angesprochen werden. Die für experimentelles Design erforderliche Bildung von tatsächlich äquivalenten Vergleichsgruppen ist nur in wenigen Evaluationsvorhaben in der Grundlagenforschung möglich. Deshalb sind Evaluationsstudien - streng genommen - nur in der Lage, Hinweise auf Schwachstellen zu geben, nicht jedoch kausale Rückschlüsse zu liefern, wie das in experimentellen Wissenschaften der Fall ist. Anstelle des Experiments wird in der pädagogischen Evaluationsforschung mit Korrelationsstudien gearbeitet. Um zur korrelativen Aufklärung des Bedingungsgefüges beizutragen, sollten mögliche fördernde oder hemmende Aspekte eines Lernangebots (zum Beispiel Lernkultur im Unternehmens oder Bedingungen für eine erfolgreiche Implementation) in den Untersuchungsplan aufgenommen werden. Dem Evaluator sollte dabei bewusst sein, dass die von der Grundlagenforschung geprägten Idealvorstellungen bezüglich des Designs meist nur bei summativen Evaluationen erfüllt werden können.

Evaluation bedient sich sozialwissenschaftlicher Verfahrensweisen. Sie ist umfassender als einfaches „Messen“ und beinhaltet sowohl qualitative als auch quantitative Aspekte zur Beschreibung von Bildungsangeboten. Sie dient als Planungs- und Entscheidungshilfe und ist somit ziel- und zweckorientiert. Evaluation kann nur in interdisziplinärer Herangehensweise erfolgreich durchgeführt werden. Die Frage, was und wie evaluiert werden soll, lässt sich nicht generell sondern nur unter Berücksichtigung der Zielsetzung der Evaluation des jeweiligen Projekts und der zur Verfügung stehenden Ressourcen beantworten.

Literatur

- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin/Heidelberg: Springer.
- Dickhäuser, O. (2001). *Computernutzung und Geschlecht*. Münster: Waxmann.
- Dillon, A. & Gabbard, R. (1998). Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research*, 68, S. 322-349.
- Fricke, R. (2000). Qualitätsbeurteilung durch Kriterienkataloge. In P. Schenkel, S.-O. Tergan & A. Lottmann (Hrsg.). *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Nürnberg: Bw Bildung und Wissen Verlag und Software GmbH, S. 75-88.
- Gräber, W. (Hrsg.). (1990). *Das Instrument MEDA*. Kiel: IPN.
- Greve, W. & Wentura, D. (2002). *Wissenschaftliche Beobachtung*. Weinheim: Beltz-PVU.
- Günther, W. & Mandl, H. (Hrsg.). (1997). *Telelearning: Aufgabe und Chance für Bildung und Gesellschaft*. Bonn: Deutsche Telekom.
- Gunn, C. (2002): *System Log Data*. Verfügbar auf http://www.icbl.hw.ac.uk/lti/cookbook/system_log_data/indx.html#enhead [Stand Dezember 2002].
- Hammann, P. & Erichson, B. (2000). *Marktforschung*, 4. Aufl. Stuttgart: Lucius & Lucius.
- Heckhausen, H. (1989). *Motivation und Handeln*. Berlin: Springer.
- Kim, H.O. & Klauser, F. (2004). Analyse der Erfolgsfaktoren für computer- und netzbasierte Lernangebote. (Research Report Impuls^{EC} 8). Universität Osnabrück.
- Klauser, F. & Kim, H.-O. (2003a). Zielgruppenanalyse – Grundlage für die effektive Entwicklung und Implementation netzbasierter Lernumgebungen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 99,1, S. 26-41.

- Klauser, F. & Kim, H.-O. (2003b). Analyse der Zielgruppen für IMPULS^{EC}: Konzeption, Befunde und Diskussion. (Research Report Impuls^{EC} 5). Universität Osnabrück.
- Klauser, F., Kim, H.-O. & Pollmer, M. (2003). Produktqualität und individuelle Leistungsvoraussetzungen – Determinanten für den effektiven Einsatz von E-Learning-Angeboten in der Berufsausbildung. *Erziehungswissenschaft und Beruf*, 51,1, S. 29-56.
- Lazarus, R.S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Lemke, S.G. (1995). *Transfermanagement*. Göttingen: Verlag für Angewandte Psychologie.
- Lienert, G. A. & Raatz, U.(1994). *Testaufbau und Testanalyse*, 5. Aufl. Weinheim: Beltz.
- Mandl, H. & Reinmann-Rothmeier G. (2000). Von Qualitätsbewusstsein über Selbstevaluation und maßgeschneidertes Vorgehen zur Transfer-sicherung. In P. Schenkel, S.-O. Tergan & A. Lottmann (Hrsg.). *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Nürnberg: BW Bildung und Wissen, S. 89-105.
- Meier, A. (2000). MEDA und AKAB: Zwei Kriterienkataloge auf dem Prüfstand. In P. Schenkel, S.-O. Tergan & A. Lottmann (Hrsg.). *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Nürnberg: Bw Bildung und Wissen Verlag und Software GmbH, S. 164-189.
- Mogey, N. (Hrsg.). *Learning Technology Dissemination Initiative: Evaluation Studies*. Verfügbar auf http://www.icbl.hw.ac.uk/ltidi/evalstudies/es_all.pdf. [Stand: Januar 2003].
- Pekrun, R. & Hofmann, H. (1999). Lern- und Leistungsmotivationen: Erste Befunde eines Forschungsprogramms. In M. Jerusalem & R. Pekrun (Hrsg.). *Emotion, Motivation und Leistung*. Göttingen: Hogrefe, S. 247-268.

- Prenzel, M. & Drechsel, B. (1996). Ein Jahr kaufmännische Erstausbildung: Veränderungen in Lernmotivation und Interesse. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 92, S. 327-344.
- Rheinberg, F. (1997). *Motivation*. Stuttgart: Kohlhammer.
- Rowntree, D. (1992). *Exploring open and distance learning*. London: Kogan Page.
- Sächsisches Ministerium für Kultus (2002). *Bewertung, Einsatz und Evaluation von Lernsoftware im Englischunterricht der Berufsschule im Berufsfeld Wirtschaft und Verwaltung*. Dresden.
- Scriven, M. S. (1972). Die Methodologie der Evaluation. In Ch. Wulf (Hrsg.). *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München: Piper, S. 60-91.
- Scheele, B. & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien*. Tübingen: Francke Verlag.
- Schweer, M. K. W. (1998). Evaluation der Lehre. In D. H. Rost (Hrsg.). *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz PVU, S. 122-125.
- Simon, B. (2001). *E-Learning an Hochschulen: Gestaltungsräume und Erfolgsfaktoren von Wissensmedien*. Köln: Josef Eul Verlag.
- Sonntag, K. (1996). *Lernen im Unternehmen. Effiziente Organisation durch Lernkultur*. München: C. H. Beck.
- Tergan, S.-O. (2000). Grundlagen der Evaluation: ein Überblick. In: P. Schenkel, S.-O. Tergan & A. Lottmann (Hrsg.). *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Nürnberg: BW Bildung und Wissen, S. 22-51.
- Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & B. Weidenmann (Hrsg.). *Pädagogische Psychologie*, 4. Auflage. Weinheim: Beltz-PVU, S. 415-465.
- Wild, K-P. (2000). *Lernstrategien im Studium*. Münster: Waxmann.
- Wottawa, H (2001). Evaluation. In A. Krapp & B. Weidenmann. (Hrsg.). *Pädagogische Psychologie*. Weinheim: Beltz PVU, S. 647-674.
- Wottawa, H. & Thierau, H. (1998). *Lehrbuch Evaluation*. Bern: Huber.

- Wulf, C. (1972). *Evaluation: Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München: Piper.
- Wulf, C. (1975). Funktionen und Paradigmen der Evaluation. In K. Frey (Hrsg.). *Curriculum-Handbuch*. München: Piper, S. 580-600.

ISBN 3-936475-13-X