

# Evaluationsverfahren für den Vergleich virtueller Lernplattformen

*Baumgartner, Peter, Hartmut Häfele und Kornelia Maier-Häfele (2003).  
Evaluationsverfahren für den Vergleich virtueller Lernplattformen. In: Online lehren. Planung  
und Gestaltung netzbasierter Weiterbildung. Hg: Heino Apel und Susanne Kraft. Bielefeld,  
Bertelsmann. S. 219-236.*

## **Abstract:**

Im Auftrag des österreichischen Bundesministeriums für Bildung, Wissenschaft und Kultur (bm:bwk) haben wir einen weltweiten Vergleich von Lernplattformen durchgeführt. Nachfolgender Beitrag beschreibt zuerst das allgemeine methodische Konzept, den konkreten Ablauf des Verfahrens und fasst im letzten Abschnitt die Ergebnisse dieses Evaluierungsprozesses zusammen.

Als Methode für das Evaluierungsverfahren wurde die Qualitative Gewichtung und Summierung (QGS) gewählt. Die Begründung dieser Methode wurde bereits am Beispiel zweier Preisausschreiben - EASA [W1] und MeDiDa-Prix [W2] – sowohl theoretisch als auch in ihrer praktischen Handhabung ausführlich beschrieben (Baumgartner, Payr, 1997). Da das Verständnis der Methode für die Diskussion der Ergebnisse des von uns durchgeführten weltweiten Vergleichs von Lernplattformen entscheidend ist, stellen wir das theoretische Konzept hier nochmals vor.

## **1. Das Evaluationsverfahren (Theorie und Konzept)**

Es ist Standard, dass bei vergleichenden (Produkt-)Untersuchungen Kriterienkataloge eingesetzt werden. In Form von Prüf- oder Checklisten soll eine Bewertung der Qualität der zu evaluierenden Produkte erfolgen. Für die große Verbreitung dieses Verfahrens sind mehrere Vorteile verantwortlich:

⊇ *Kostengünstig*: Es genügt eine fachkundige Person, ein Testaccount bzw. eine Programmkopie und eine Hardwareausstattung, auf der die Software mit all ihren Optionen lauffähig ist.

⊄ *Schnell*: Durch die Trennung des Medienproduktes von seinem realen Einsatzgebiet (Lehrgeschehen) kann die Bewertung rasch erfolgen.

⊂ *Methodisch sauber*: Durch die schrittweise, immer gleiche Abarbeitung umfangreicher Kriterienlisten erscheint das Verfahren objektiv und methodisch sauber.

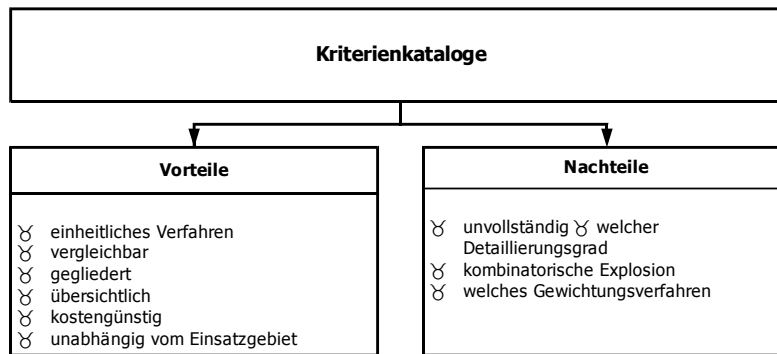


Abb. 1: Vor- und Nachteile von Kriterienkatalogen

So wichtig und verdienstvoll die Bewertung mittels Kriterienkatalogen auch sein mag, so weist diese Vorgehensweise auch einige gravierende methodische Mängel auf (Baumgartner, Häfele, Maier-Häfele 2002b), von denen hier das Gewichtungsproblem herausgegriffen werden soll:

Obwohl über die Items eines Kriterienkatalogs relativ schnell Einigkeit geschaffen werden kann, ist die Frage, wie nun die einzelnen Faktoren gewichtet werden sollen, meist sehr umstritten. Gerade die Gliederung und Gewichtung der einzelnen Kriterien entscheidet aber über die vergleichende Bewertung und Auswahlentscheidung.

Es gibt zwei Verfahren zur relativen Bewertung einzelner Items eines Kriterienkatalogs - numerische und qualitative Gewichtungsprozeduren:

### 1.1. Die Numerische Gewichtung und Summierung (NGS)

Numerisches Gewichten und Summieren (NGS) stellt derzeit das dominante Modell für komplexe Produktevaluationen dar und wird insbesondere im Zusammenhang mit der Bewertung von Lernsoftware in Form von Check- oder Prüflisten angewendet (vgl. Thomé 1988). Das NGS-Verfahren kommt in verschiedenen Formen vor und kann sowohl beschreibend als auch bewertend eingesetzt werden. Die Prozedur dabei ist folgende:

- Zuerst werden die einzelnen Dimensionen in ihrer relativen Wertigkeit (z. B. anhand einer 1 bis 3-, 1 bis 5- oder 1 bis 7-Skala) eingeschätzt bzw. gewichtet.
- Anschließend wird die Leistung des Evaluanden [= die zu evaluierende Sache (z.B. Lernsoftware)] nach den einzelnen Dimensionen eingeschätzt (=rating).
- Das Produkt von Leistungsbewertung und Gewicht (Leistungspunkte x Gewichtung) wird berechnet und für jeden einzelnen Evaluanden summiert.
- Es ergibt sich für jeden Evaluanden eine einzige Zahl, die den relativen Rang des jeweiligen Evaluanden bestimmt. Sieger ist der Evaluand mit der höchsten Punktezahl.

Das NGS-Verfahren ist infolge einer Reihe von Vorteilen (es ist leicht verständlich, einfach durchzuführen, aufschlussreich und ergibt manchmal auch valide Ergebnisse) sehr beliebt. Obwohl es immer einen ersten Aufschluss bzw. Einblick bietet und daher im Rahmen einer weiterführenden Evaluation durchaus brauchbar ist, hat dieses Verfahren schwerwiegende methodische Mängel, so dass die alleinige Verwendung dieses Verfahrens ungeeignet ist:

*Das Minimum-Problem:* Ein Set von Gewichten löst nicht das Problem, dass einige Dimensionen (Merkmale) erst dann eine sinnvolle Funktion des Evaluanden darstellen, wenn ein bestimmtes Mindestmaß überschritten ist. In einer abschließenden Summierung zu einer einzigen Zahl gehen diese inhaltlichen Minimalanforderungen jedoch verloren. Diese Schwierigkeit lässt sich jedoch durch eine Erweiterung des NGS-Verfahrens beheben (NGS-Modell mit Minima). Dabei wird jedes der Minima-Kriterien zuerst geprüft, bevor die Analyse fortgesetzt wird. Nur eine Leistung über dem erforderlichen Minimum wird gewichtet. Evaluanden, bei denen einzelne Kriterien dieses notwendige Minimum nicht erreichen, scheiden aus (K.O.-Kriterien).

*Das Problem des Detaillierungsgrades:* Eine wesentliche Kritik am NGS-Verfahren besteht darin, dass die Anzahl der Kriterien nicht voraussehbar ist. Sie kann von etwa einem Dutzend bis zu einigen hundert Kriterien reichen. Damit werden aber entweder wichtige Dimensionen durch eine Vielzahl von Trivialitäten überschwemmt oder aber weniger wichtige Faktoren wirken sich auf das Gesamtergebnis zu stark aus. Das Festlegen einer fixen Punkteanzahl, die nicht überschritten werden darf, reduziert zwar das Problem, kann es aber nicht gänzlich lösen. Was sind die relevanten Kriterien (Wie viele? Wie detailliert?) und vor allem: welche Gewichtung kommt ihnen jeweils zu?

Besonders fatal beim NGS-Verfahren ist es, dass diese Gewichtungsprozedur keine Spuren hinterlässt. Da sich als Ergebnis bloß eine einzige Zahl pro Evaluand ergibt, sind nachträglich keine inhaltlichen Fehlerkorrekturen mehr möglich.

*Annahme einer linearen Skala:* Die NGS-Methode nimmt eine lineare Skala der Nützlichkeit (Vergabe von Punkten und Summierung) an, was jedoch falsch ist. Die verschiedenen Komponenten des Evaluanden lassen sich nicht über eine einzige Skala bewerten.

Das ist die entscheidende Kritik und das innerhalb des NGS-Verfahrens unlösbare Hauptproblem dieser Methode. Die einzelnen Kriterien (z.B. Performance, Kommunikationswerkzeuge etc.) liegen nicht auf einer linearen Skala (z.B. des Lernerfolges). Aber nur bei metrischen Skalen (Intervall- und Ratioskala) sind Rechenoperationen wie Addition und Multiplikation erlaubt.

Die scheinbar methodisch sauberen und exakten Ergebnisse des NGS-Verfahrens sind ein methodischer Trugschluss, eine Illusion, welche nicht der Realität entspricht.

Obwohl mit dem NGS-Verfahren viele Evaluanden in einem ersten Durchgang provisorisch miteinander verglichen werden können und es als erster grober Filter durchaus brauchbar ist, ist es letztlich notwendig, einen paarweisen Vergleich mit qualitativen Bewertungsverfahren durchzuführen.

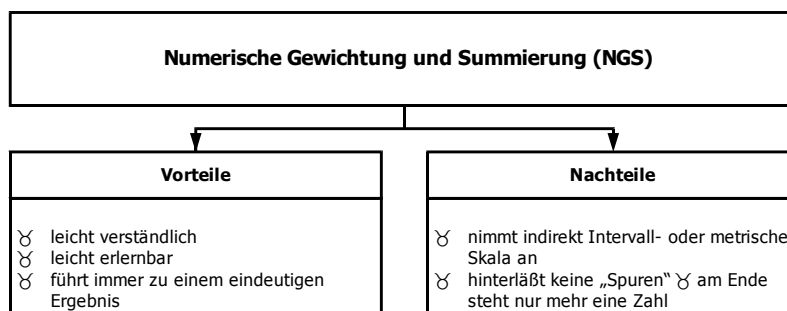


Abb. 2: Vor- und Nachteile der Numerischen Gewichtung und Summierung

## 1.2. Die Qualitative Gewichtung und Summierung (QGS)

Viele der oben unter NGS beschriebenen Nachteile lassen sich bei einer qualitativen Vorgehensweise beseitigen. Nachfolgend beschreiben wir die Schritte, um eine Qualitative Gewichtung und Summierung durchzuführen.

In einem ersten Schritt werden noch nicht die einzelnen Evaluanden betrachtet, sondern die Liste der Kriterien, welche in den Evaluationsprozess mit einbezogen werden soll. Für diese Kriterien werden nun fünf Gewichte vergeben. Es empfiehlt sich dafür nicht-ziffernartige Symbole zu verwenden, damit gleich zu Beginn eine Verwendung als Intervall- oder Ratioskala und damit Rechenoperationen ausgeschlossen werden.

Bewährt hat sich die folgende Einteilung:

- Essentiell (E)
- Äußerst Wichtig (\*)
- Sehr Wichtig (#)
- Wichtig (+)
- Weniger Wichtig (l)
- Nicht Wichtig (0)

Damit werden einmal die Gewichte der einzelnen Merkmale festgelegt (von 0 - „Nicht Wichtig“ bis \* - „Äußerst Wichtig“), darüber hinaus wird auch bestimmt, welche Eigenschaften Minimalerfordernisse (E - „Essentiell“) darstellen.

Alle 0-Dimensionen können nun gestrichen werden. Sie sind als „Nicht Wichtig“ gewichtet worden und daher für die weitere Bewertung irrelevant. Damit wird unnötiger Analyseaufwand vermieden.

Nun werden zum ersten Mal die einzelnen Evaluanden bewertet: Es wird überprüft, ob alle Evaluanden die Minimalerfordernisse (Kriterien, die mit E gewichtet wurden) auch tatsächlich erfüllen. Falls nicht, werden sie aus der weiteren Analyse ausgeschieden. Dadurch wird wiederum der weitere Arbeitsaufwand beträchtlich reduziert. Allerdings ist dafür Sorge zu tragen, dass es sich dabei um ein alles-oder-nichts-Attribut handelt (z.B. die Lernplattform ist in Form einer nationalen Rahmenlizenz einsetzbar ja/nein).

Andernfalls muss das (Anspruchs-)Niveau, das unbedingt erforderlich ist, genau festgelegt und geprüft werden, ob der betreffende Evaluand dieses erreicht oder nicht. Nun sind bereits alle mit 0 gewichteten Kriterien und jene Evaluanden, welche die Mindestkriterien (= E) nicht erfüllen, aus dem Evaluationsprozess ausgeschieden. Die zu untersuchenden Kriterien der verbleibenden Evaluanden weisen jetzt nur mehr Unterschiede zwischen \* und I auf.

Der nächste Schritt ist die Bewertung dieser Kriterien für jeden einzelnen Evaluanden. Das heißt, dass die Erfüllung jedes Kriteriums für die einzelnen Evaluanden überprüft und bewertet wird. Die Bewertung erfolgt in einem Rahmen von 0 bis zur maximalen Gewichtung des jeweiligen Kriteriums. D. h. ein #-Kriterium, das im ersten Schritt mit # gewichtet wurde, kann keinen höheren Wert als # erhalten (also nur 0, I, +, #).

Nun werden für jeden Evaluanden die Summen der mit \*, # mit + und I gewerteten Kriterien gebildet. Dies geschieht dadurch, dass alle \*, alle #, alle + und alle I separat zusammengezählt werden.

Sie erhalten nun eine Liste etwa der folgenden Art:

Produkt A: 7 \*, 5 #, 2 +, 3 I

Produkt B: 2 \*, 10 #, 3 +, 2 I

Produkt C: 3 \*, 7 #, 1 +, 6 I

etc. ...

Nach den bisherigen Verfahrensschritten entsteht eine Rangordnung (ranking), die anschließend auch mit einer integrierenden Schlussbewertung (grading) versehen werden kann (z. B. „kommt in die TOP 15“ etc.). Diese ließe sich z. B. durch Festlegung einer Minimumanzahl von \* oder \* und # oder auf einer individuellen Fallbasis durchführen, nachdem alle Evaluanden bereits bewertet wurden.

Im nächsten Schritt werden nun jene Eigenschaften, die alle Evaluanden gleichermaßen aufweisen (z. B. wenn alle Evaluanden ein bestimmtes Kriterium mit + erfüllt haben), ausgeschlossen. Damit wird der weitere Vergleich auf einer Fall-zu-Fall-Basis vereinfacht.

Abschließend kann nun geprüft werden, ob bereits eine eindeutige Rangordnung möglich ist. Eindeutig heißt, dass z. B. ein Evaluand mit 3\*, 4# und 2+ auf jeden Fall besser ist als einer mit 2\*, 4# und 3+. Hat jedoch der zweite Evaluand z. B. 2\*, 7#, so ist keine eindeutige Entscheidung möglich und die beiden Kandidaten müssen in einem paarweisen Vergleich genauer untersucht werden.

Wenn nun nicht bereits entscheidbare Verhältnisse vorliegen, so kann eine neuerliche Gewichtung im Lichte der vergleichenden Bewertung konkreter Einzelfälle hilfreich sein.

Neben seiner relativen Komplexität hat das QGS-Verfahren den Nachteil, dass es keinen definitiven Entscheidungsalgorithmus hat. Manchmal muss es als iterative Prozedur mehrfach durchlaufen werden und im Lichte der bisherigen Analyse müssen einzelne Bewertungen nochmals durchgeführt werden. Das Verfahren wechselt damit ständig zwischen holistischer und analytischer Betrachtungsweise, ergibt aber immer sinnvolle und vor allem nachvollziehbare und überprüfbare Ergebnisse.

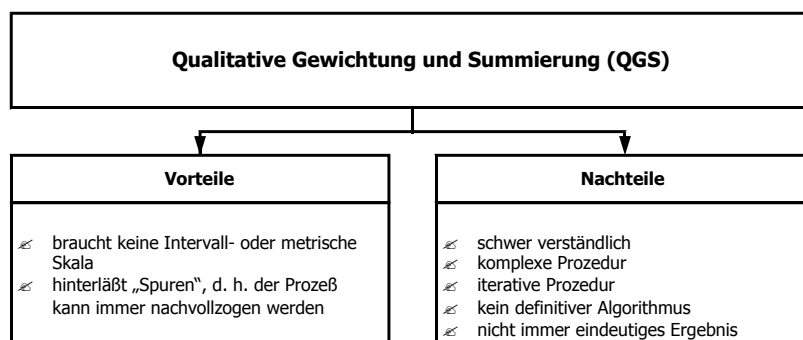


Abb. 3: Vor- und Nachteile der Qualitativen Gewichtung und Summierung

## 2. Die praktische Durchführung der Evaluation

Das im vorliegenden Fall vorgestellte Evaluierungsverfahren gliedert sich in zwei Abschnitte:

### Abschnitt 1:

- Weltweiter Vergleich aller relevanten Lernplattformen.
- Erstellung einer Liste der Top 15.
- Qualitative Beschreibung dieser 15 Produkte.

### Abschnitt 2:

- Detaillierter Vergleich der Produkte aus der Top-Liste.
- Erstellung einer Liste der Top 8.
- Installation dieser 8 Produkte und Durchführung eines Usability-Tests
  - durch das Evaluationsteam,
  - durch das Einholen von Referenzen,
  - durch Befragung und Beobachtung von Anwender(inne)n der potentiellen Zielgruppen.

## 2.1 Abschnitt 1

Der erste Teil des Verfahrens selbst wiederum lässt sich in drei verschiedene Abschnitte unterteilen:

- Erarbeiten der Evaluierungskriterien.  
Definition der Mindestkriterien (K.O.-Kriterien, Essentials) und qualitative Gewichtung der Kriterien. Parallel dazu erfolgt die Sichtung und Dokumentation der am Markt befindlichen Lernplattformen.
- Anwenden des Kriterienkataloges und der Mindestanforderungen.
- Erstellung der Shortlist und Beschreibung der 15 Top-Produkte.

Darüber hinaus haben wir auch die Verfahren der „Expert(inn)engruppen“ und das Review-Verfahren in die Prozedur mit einbezogen.

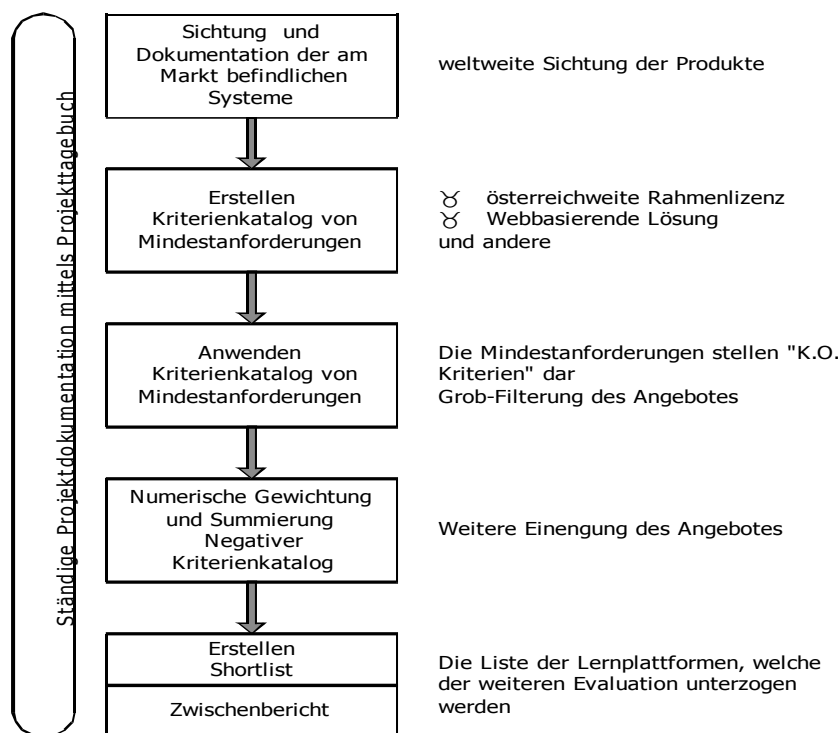


Abb. 4: Projektdesign der ersten Evaluationsphase

### 2.1.1 Die Ausarbeitung der Evaluierungskriterien

#### a) Aufbau von Websites

Unser erster Schritt im Verfahren war es, Websites aufzubauen, auf denen wir vom Fortgang unserer Untersuchung kontinuierlich berichteten. Es gab dabei sowohl eine Website zum Gesamtprojekt und unseren detaillierten

Erfahrungen (Projekttagbuch) als auch eine Community-Website [W3], auf der wir die Untersuchungsergebnisse dokumentierten.

Im Zusammenhang mit dem zweiten Schritt, dem Aufbau einer Expert(inn)engruppe, war dies für einen geordneten und ungestörten Verlauf des Evaluierungsverfahrens sehr wichtig. Es gab relativ wenige Interventionen seitens unterschiedlicher Stakeholder (z.B. den Herstellerfirmen von Lernplattformen) oder von Verantwortlichen für e-Learning in Organisationen, die bereits auf eine bestimmte Lernplattform „eingeschworen“ waren.

### **b) Das Expert(inn)enteam**

Unser zweiter vorbereitender Schritt im Verfahren war es, gemeinsam mit dem Auftraggeber eine Expert(inn)engruppe aufzubauen, mit der sowohl die Art der Kriterien, als auch ihre relative Gewichtung zueinander abgestimmt wurden. Nachträglich gesehen war dies ein zentraler Punkt nicht nur für die Akzeptanz der Ergebnisse sondern auch für das Gelingen des insgesamt sehr heiklen – von verschiedenen Interessenorientierungen durchdrungenen – Verfahrens.

In zwei von uns moderierten Sitzungen hat diese Gruppe sowohl die Kriterienliste (inklusive K.O.-Kriterien) erstellt, als auch die Gewichtung der Kriterien vorgenommen. Zur Vorbereitung dazu gab es ein einleitendes Referat über Funktion und Wirkungsweise von Lernplattformen.

Zur gemeinsamen Ausarbeitung der Kriterien haben wir die Delphi-Methode (Seeger 1979) gewählt, d. h. wir haben nach der Sammlung und Dokumentation der eingereichten Stellungnahmen die Ergebnisse der gesamten Gruppe rückgemeldet und sie um Streichungen und Ergänzungen gebeten. Diese Vorschläge wurden wiederum in die Kriterienliste eingearbeitet und an die Experten rückgemeldet.

### **c) Die Mindestkriterien**

Die Mindestkriterien dienten vor allem dazu, eine grobe Sichtung der am Markt befindlichen Systeme anzuleiten und reine Präsentations-, Verwaltungs- und Authoring-Tools von Lernmanagement-Systemen abzugrenzen. Wichtig war uns, dass sowohl didaktische Funktionen des interaktiven Online-Lernens als auch administrative Funktionen integriert waren.

Die Mindestkriterien wurden als sogenannte K. O.-Kriterien angewendet, d.h. jene Systeme, die diese nicht erfüllten, wurden bei der nächsten Evaluationsphase nicht mehr berücksichtigt.

Dazu wurden die Anbieter von Lernplattformen angeschrieben. Lieferten die Hersteller trotz Nachfrage und einer angemessenen Nachfrist keine Informationen, schieden die Produkte aus.

Die folgenden Kriterien wurden als Mindestkriterien (K. O.-Kriterien) angewendet:

- **Bundesweite Rahmenlizenz:**  
Es muss eine bundesweite Rahmenlizenz angeboten werden. Ausgeschlossen wurde damit ein Pay-Per-User-Lizenzmodell.
- **Webbasierte Lösung:**  
die auf einem Standardprotokoll (TCP/IP) und einem Standardwebbrowser beruht. Die Anwender(innen) können mittels beliebigem Webbrowser am Lernprozess teilnehmen. Das Lernmanagement sowie das Ergebnis einer Content-Generierung mittels Wizards bzw. Templates muss die Möglichkeit enthalten, mit einem Webbrowser ohne jegliche Erweiterung (Viewer und Plugins) bedienbar zu sein.
- **Administration:**  
Die Verwaltung von Userinnen sowie die Organisation (Erstellen, Erweitern, Löschen) von Content (Lerninhalte) und Kursen muss möglich sein.  
Das System muss die Authentisierung und Autorisierung (Rollen, Gruppen, Rechte) vom entsprechenden Zugriffs- und Rechtemanagement übernehmen können. Konkret muss das System LDAP-fähig (Light Weight Directory Access Protocol) oder entsprechend anpassbar sein.  
Das System muss rollenbasierend sein.  
Das System muss multilingual sein und zumindest die deutsche und englische Sprache einschließen.
- **Kommunikation:**  
innerhalb der Rollen (Lernende, Lehrende) und zwischen den Rollen (Lernende-Lehrende) muss in elektronischer Form unterstützt werden.
- **käuflich erwerbbar / open source / public domain / LSP-Lösung:**  
Das System sollte käuflich erwerbbar, open source, public domain oder eine Mietlösung sein.  
Einige Anbieter im e-Learning Bereich bieten Lernplattformen nur auf Mietbasis auf deren eigenen Servern an (LSP - Learning Service Provider). Eine LSP-Lösung wird für den post-sekundären und tertiären Bereich für möglich gehalten, für den Schulbereich jedoch ausgeschlossen.

Alle Anbieter von e-Learning-Plattformen, auf die wir im Rahmen unserer Recherche gestoßen sind, wurden angeschrieben und von der Evaluation informiert. In diesem ersten Schreiben haben wir den Auftraggeber der Evaluation und die Evaluationsziele bekannt gegeben, kurz das Projektdesign skizziert und die Unternehmen aufgefordert, zu den Mindestkriterien Stellung zu nehmen.

Gleichzeitig haben wir auch Web- und Literaturrecherche betrieben und die bisherigen Erfahrungen der User als erste Ansatzpunkte in die Dokumentation einbezogen. Es hat sich allerdings gezeigt, dass die Rezensionen (Reviews) entweder stark subjektiv (individuelle Erfahrungsberichte) waren oder sich auf einige willkürlich herausgegriffene Plattformen beschränkten. Auch in den wenigen Fällen, in denen es umfangreichere Listen und Darstellungen gab, fehlte es an einer begründeten Gewichtung.

Unsere Community-Website [W3] diente ab diesem Projektabschnitt als Informationsplattform, auf der wir die Mindestkriterien sowie weitere Informationen zum Projekt zur Verfügung stellten.

Der nächste Schritt im der ersten Phase war es, mit den nach der Anwendung der Mindestkriterien verbliebenen 40 Lernplattformen die Qualitative Gewichtung und Summierung anhand eines Kriterienkataloges vorzunehmen.

#### **d) Die Kriterien der Qualitativen Gewichtung und Summierung (QGS)**

Um für die zweite Sitzung mit den Expert(inn)en eine Kriterienliste für die Qualitative Gewichtung und Summierung (QGS) vorbereiten zu können, haben wir bereits in der ersten Sitzung eine Liste mit „Dimensionen zur Beschreibung der Kriterien für die QGS“ vorgestellt. Diese Dimensionen, die noch keine Kriterien, sondern eher Überbegriffe der späteren Kriterien darstellten, wurden hinterfragt, diskutiert, ergänzt und gestrichen. So wurde uns klar, in welche Richtung die Wünsche der Expert(inn)en und damit auch jene unserer Auftraggeber gingen.

Diese Vorgehensweise ermöglichte es uns, bereits zum zweiten Zusammentreffen eine Liste von Kriterien zur Qualitativen Gewichtung und Summierung mitzubringen, die nur noch eine Feinabstimmung benötigte. So war es möglich, in nur zwei Sitzungen auch die Gewichtung der Kriterien für die QGS durchzuführen.

Die Kriterien für die QGS lassen sich unter den folgenden Kategorien zusammenfassen:

- Kommunikation, Kooperation und Kollaboration (5 Kriterien)
- Didaktik (7 Kriterien)
- Administration (3 Kriterien)
- Technik (12 Kriterien)

Unser Hauptproblem war es, den Kriterien aus dem pädagogischen Bereich (Didaktik und Kommunikation) eine größere Gewichtung im Gesamtverfahren zukommen zu lassen. In der ersten Fassung gab es im Bereich der Technik und Administration viele Kriterien, die einen sehr feinen Detaillierungsgrad hatten. Dadurch bestand die Gefahr, dass diese Merkmale alleine durch ihre Menge das Evaluierungsverfahren dominierten. Nachdem wir einerseits im Technikbereich ähnliche Funktionen zusammengefasst und andererseits die pädagogischen Kriterien stärker unterteilt hatten, erhielten wir ein relativ

ausgeglichenes Gesamtbild: 12 didaktisch-kommunikative zu 15 technisch-administrativen Kriterien.

Diese 27 Kriterien wurden der Expert(inn)engruppe zur individuellen Gewichtung vorgelegt. Die Aufgabe der Expert(inn)en bestand darin, jedes Kriterium mit einem der nachfolgenden Symbole zu bewerten. Die Symbole wurden verwendet, um von vornherein eine Addition der Merkmale zu einer Gesamtzahl zu verhindern.

Es ergab sich die folgende Gewichtung (gewichtete Kriterienliste):

### Legende:

*	=	äußerst wichtig
#	=	sehr wichtig
+	=	wichtig
	=	weniger wichtig
0	=	nicht wichtig

### Kommunikation, Kooperation und Kollaboration:

synchrone Kommunikationstools (Chat, Whiteboard, etc.)	*
asynchrone Kommunikationstools (E-Mail, Foren, etc.)	*
Annotationen (Anmerkungen zum Content)	#
Gruppenbildung durch Rollen (Lehrende, Lernende Tutoren)	*
externe Kommunikationstools integrierbar	#

### Didaktik:

lässt verschiedene Lehr- und Lernmodelle zu	*
durchführen von Tests möglich	#
interaktive Übungen einzeln und im Team möglich	#
Modularisierung von Lerninhalten möglich (Learning Objects)	*
Feedback zum Lernfortschritt (Statistiken, Terminkontrollen)	#
Autorenfunktionen integriert (benutzerfreundliche Vorlagen)	+
Learning-flow-Management (Lernpfade)	#

### Administration:

Tracking von Usern (Protokollierung der Aktivitäten)	#
Rechnungsverwaltung (ERP-Schnittstelle vorhanden)	+
Personalisierung möglich	*

### Technik:

Anpassbarkeit des Systems (CI übernehmbar, etc.)	*
--	---

Erweiterbarkeit (Schnittstellen, Makros, etc.)	*
Skalierbarkeit	*
Distribuierbarkeit (System kann verteilt betrieben werden)	*
entspricht aktuellen e-Learning Standards (AICC, SCORM, etc.)	*
unterstützt Open-Source Serverbetriebssysteme	#
Support (Response-Zeit, Erreichbarkeit, Sprache, etc.)	*
Dokumentation (Referenz-Handbücher, Tutorials, etc.)	#
Sicherheit bei Datenweiterreichung	*
Ressourcenbedarf	#
österreichischer Vertriebspartner vorhanden	#
unterstützt Standardobjekttypen und -formate (PDF, etc.)	*

Es fällt auf, dass die 0 (nicht wichtig) nie vergeben wurde. Da es dieselben Personen waren, die die Kriterienliste erstellten und anschließend gewichteten, ist dies nur folgerichtig. Funktionen, die als unbedeutend eingeschätzt worden sind, sind gar nicht erst als Kriterien aufgestellt worden.

## 2.1.2 Die Anwendung der Evaluierungskriterien

Im Anschluss an die Stellungnahme zu den Mindestkriterien haben wir die 40 noch in der Evaluation verbliebenen Unternehmen aufgefordert, zu den Kriterien für die Qualitative Gewichtung und Summierung Stellung zu nehmen. Die Liste der Kriterien sowie die Hinweise darüber, wie die Kriterien gewichtet werden, wurden wiederum auf der Community-Webseite [W3] zur Verfügung gestellt.

Sehr viel Augenmerk haben wir auf eine präzise Ausformulierung der Stellungnahmen zu den QGS verwendet. Dies stellte für uns einen erheblichen Aufwand dar, da alle Stellungnahmen durchgearbeitet und jedem Unternehmen erneut explizite Fragen zu den gegebenen Antworten aus der ersten Befragung gestellt werden mussten. Dieser Aufwand ermöglichte es uns jedoch, die Bewertung der Kriterien für die QGS der einzelnen Produkte so fair wie möglich zu gestalten.

Auch hatten wir alle Verantwortlichen um die Zusendung von Whitepapers zum Produkt gebeten, um möglichst viele aussagekräftige Zusatzinformationen zu erhalten. Dieser Bitte kamen alle Unternehmen nach, allerdings nur nach der ausdrücklichen Zusage unsererseits, diese ausführlichen Produktinformationen nicht an andere Firmen weiterzugeben („non disclosure agreement“). Insgesamt konnten so ca. 3.100 Seiten an technischen Informationen zusammengetragen werden (94 Megabyte an PDF- und Word-Dokumenten).

Im Projektverlauf stellte sich heraus, dass wir eine zusätzliche Evaluationsschleife einbauen mussten, um so gründlich wie möglich zu bewerten. Die in der Theorie vorgesehene Bewertung der Kriterien durch Fragebögen, Telefonate etc. reichte nicht aus. Gerade bei den didaktischen

Funktionen fehlten den Befragten entweder das Know-how oder diese Kriterien waren nicht eindeutig beantwortbar.

So forderten wir von allen Herstellern, deren Produkte die Mindestkriterien erfüllten, einen Testaccount an. Diese Bitte stellte einige Unternehmensrepräsentanten vor große Schwierigkeiten. Gerade die Qualitätssicherungssysteme der großen US-amerikanischen Firmen lassen es nicht zu, dass potentielle Käufer(innen) (oder Evaluator(inn)en) das Produkt zu Gesicht bekommen, ohne dass ein(e) Repräsentant(in) der Firma zugegen ist. Dies hätte bedeutet, dass wir den freundlichen Einladungen zur Produktpräsentation quer durch Europa hätten Folge leisten müssen, oder, dass Verkäufer(innen) die Produkte bei uns vor Ort präsentiert hätten. Beidem konnten wir nicht zustimmen, da dies eine Verzerrung zugunsten jener, die das Produkt persönlich präsentieren können, ergeben hätte. Nach langer Überzeugungsarbeit von Europa nach Amerika (zuerst von unserer Seite aus, dann auch mit Hilfe der europäischen Niederlassungen) schafften es dann doch alle Unternehmen, einen solchen Account zur Verfügung zu stellen.

### **2.1.3 Die Erstellung der Shortlist**

Aus den auf dem Markt befindlichen Angeboten wurden durch die oben dargestellte Vorgehensweise 15 Produkte ausgewählt und zu einer Shortlist zusammengefasst. Innerhalb dieser Liste wurde jedoch keine Reihung vorgenommen.

Unser Verzicht auf ein detailliertes Ranking hat mehrere Gründe:

- Das QGS-Verfahren liefert nicht immer eindeutige Ergebnisse. Zwar lässt sich sehr gut eine Gruppierung vornehmen, ein definitives Ranking zwischen einzelnen Produkten ist aber in einigen Fällen nur durch einen intensiven Vergleich der beiden Produkte möglich und auch dann nicht immer schlüssig.
- Es wurde uns seitens einiger Firmen untersagt, ein definitives Ranking zu veröffentlichen. Zwar dürfen wir darüber informieren, ob das betreffende Produkt die Top 15 erreicht oder nicht erreicht hat; welchen Rang es dabei aber insgesamt einnimmt, diese differenzierte Aussage dürfen wir nicht publizieren.

### **2.1.4 Die Ergebnisse**

Von etwa 130 Produkten, die sich bei unserer ersten internetbasierten Recherche selbst als Lernplattformen bezeichneten und daher unsere Ausgangsmenge von zu untersuchenden Produkten darstellten, konnten bei einer ersten Sichtung entsprechend folgenden Definition nur 90 tatsächlich als Lernplattform bezeichnet werden.

*Unter einer webbasierten Lernplattform ist eine serverseitig installierte Software zu verstehen, die beliebige Lerninhalte über das Internet zu vermitteln hilft und die Organisation der dabei notwendigen Lernprozesse unterstützt.*

Diese 90 Produkte bildeten die Grundgesamtheit für unsere weitere Untersuchung. Davon haben rund 40 Produkte die Mindestkriterien erfüllt. Diese 40 Produkte haben wir der Qualitativen Gewichtung und Summierung unterzogen. Nachdem wir auch besonders die didaktischen Kriterien mit Hilfe der Testinstallationen noch genauer evaluiert hatten, konnten wir die Shortlist jener 15 Produkte erstellen, welche die Kriterien am besten entsprechend der Gewichtung des Expert(inn)enteams erfüllen.

### **Shortlist:**

<b>Produkt</b> (in alphabetischer Reihenfolge)	<b>Hersteller</b>
Blackboard	Blackboard
Clix	imc G.m.b.H.
Distance Learning System (DLS)	ets G.m.b.H.
Docent Enterprise	Docent inc.
Enterprise Learning Platform (vormals Sun Learntone)	Sun Microsystems
eLearning Suite (eLS)	Hyperwave AG
IBT Server	Time4you G.m.b.H.
iLearning	Oracle
ILF	M.I.T. newssystems G.m.b.H.
ILIAS (Open Source)	Universität Köln
Learning Space	Lotus (IBM)
Saba Learning Enterprise	Saba
Sitos Cubix	Bitmedia
Top Class	WBT Systems
WebCT	WebCT

Nachstehende Lernplattformen würden ebenfalls in der Shortlist vertreten sein, denn sowohl die Mindestkriterien als auch jene der QGS werden gut erfüllt. Jedoch sind diese Produkte nicht deutschsprachig und konnten deshalb in der von uns durchgeführten Studie nicht weiter berücksichtigt werden.

<b>Produkt</b> (in alphabetischer Reihenfolge)	<b>Hersteller</b>
eCollege	eCollege
Nebo	Learnframe

Genauere Informationen zu den Listen (Ausgangsliste, Liste der Produkte, welche die Mindestkriterien erfüllten etc.) und auch zu den Kriterien der Qualitativen Gewichtung und Summierung finden Sie auf unserer Community-Website [W3].

Die Produkte der Shortlist werden in unserer Publikation „E-Learning Praxishandbuch. Auswahl von Lernplattformen: Marktübersicht - Funktionen – Fachbegriffe“ (StudienVerlag) ausführlich beschrieben und in ihrer Funktionalität mit Bildschirmausdrucken („Screenshots“) dokumentiert.

## **2.2. Phase 2**

Die zweite Phase zeichnet sich vor allem durch den verstärkten Praxisbezug aus. Diesen stellen wir (wie oben beschrieben) durch ein mehrstufiges Vorgehen sicher.

- Praxistest der Produkte aus der „Top-15“-Liste durch das Evaluationsteam
- Usability-Testung
  - Befragung von Referenzpartner(inne)n
  - Befragung von Anwender(inne)n

Das Ergebnis ist eine detaillierte Beschreibung der Anwendungserfordernisse und von Nutzungserfahrungen aus dem Praxiseinsatz.

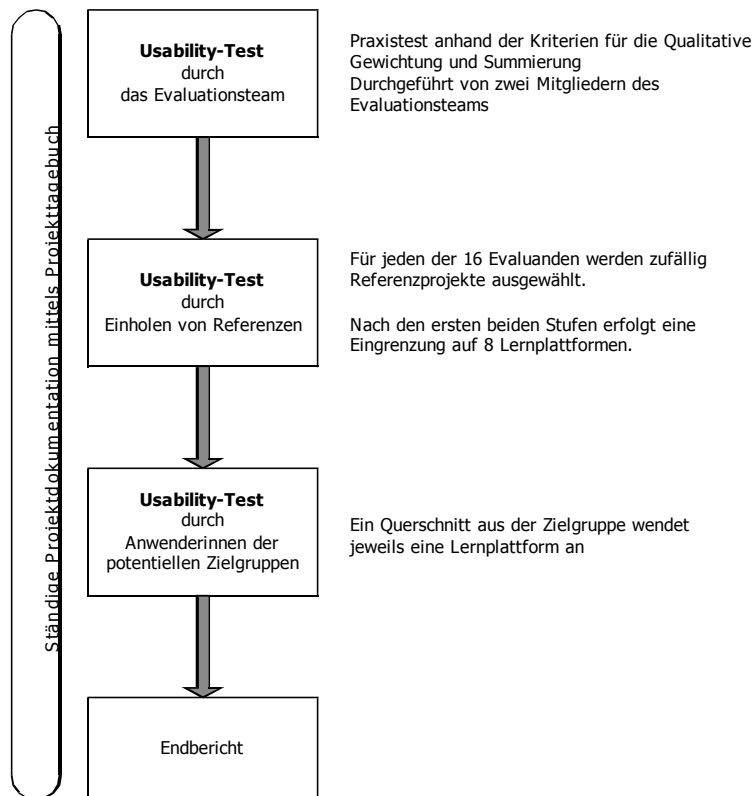


Abb. 5: Ablauf der zweiten Projektphase

## 2.2.1. Detaillierter Vergleich der Produkte –

### Erstellung der Liste der Top 8

Die untersuchten Systeme werden von mindestens zwei Mitgliedern des Evaluationsteams einem Praxistest unterzogen, der sich auf die Kriterienliste der Qualitativen Gewichtung und Summierung stützt (= Usability-Testung durch das Evaluationsteam).

## 2.2.2 Befragung von Referenzpartner(inne)n

Wir haben die Hersteller aller Produkte, welche in der Shortlist vertreten sind, um die Nennung von Referenzpartner(inne)n (Schwerpunkt Europa) ersucht. Aus den genannten Partner(inne)n haben wir zufällig jeweils drei ausgewählt und per Online-Fragebogen und Telefoninterviews zu ihren Erfahrungen mit den Plattformen befragt.

Der Rücklauf des Online-Fragebogens war mit 85 % erfreulich hoch.

## 2.2.3 Erste Ergebnisse

Die folgenden Produkte weisen in allen oben genannten Evaluationsphasen hohe Bewertungen auf (die Kreuze markieren den durch die Befragung der Referenzpartner(innen) erhobenen Einsatzbereich):

Produkte (alphabetisch sortiert)	Hersteller	Primärer Bereich	Sekundärer Bereich	Tertiärer Bereich	Erwachsenen- bildung	Corporate Learning
Blackboard	Blackboard			X		
Clix	imc G.m.b.H		X	X	X	X
DLS	ets G.m.b.H				X	
IBT Server	Time4you G.m.b.H	X	X	X	X	X
Ilias (Open Source)	Universität Köln		X	X	X	
Saba Learning Enterprise	Saba Software Inc.			X	X	X
TopClass	WBT Systems		X	X	X	X
WebCT	WebCT			X	X	

The screenshot shows the ILIAS user interface for a user named Elias Häfele. At the top, there is a navigation bar with icons for user profile, documents, folders, search, help, and other functions. Below the navigation bar, the user's name and the title 'Persönlicher Schreibtisch' are displayed. A row of links provides access to 'Persönliche Daten', 'Passwort ändern', 'Sprache ändern', 'Benutzervereinbarung', 'Wer ist online?', and 'Chat'. The main content area is divided into three sections: 'Zuletzt besuchte Lerneinheiten' (Recently visited learning units), 'Gruppenforen mit neuen Beiträgen' (Group forums with new contributions), and 'Foren von Lerneinheiten mit neuen Beiträgen' (Forums of learning units with new contributions). Each section contains a table with columns for 'Zeit' (Time), 'Lerneinheit' (Learning unit), 'Seite' (Page), 'Titel' (Title), 'Eigentümer' (Owner), and 'letzte Änderung' (Last change).

Persönlicher Schreibtisch von Elias Häfele					
Persönliche Daten	Passwort ändern	Sprache ändern	Benutzervereinbarung	Wer ist online?	Chat
Zuletzt besuchte Lerneinheiten					
Zeit	Lerneinheit	Seite			
21.01.2003	Der Abakus	Der Abakus			
20.01.2003	Psychologie der Personalbeurteilung	Startseite			
Gruppenforen mit neuen Beiträgen					
Titel					letzte Änderung
Diskussion					26.Nov.02 21:54
Foren von Lerneinheiten mit neuen Beiträgen					
Titel	Eigentümer	letzte Änderung			
Psychologie der Personalbeurteilung		29.Jan.03 10:36			

Abb. 6: Das Open Source Learning Management System ILIAS (www.ilias.uni-koeln.de).

**LERN EFFEKT** MYWEBCT | KURS WIEDERAUFNEHMEN | KURS-PLAN | QUELLEN | HILFE

Navigation verbergen **Ein Test zu Ihrer Menschenkenntnis:** Ansicht Designer-Optionen  
**Testkurs Universität Innsbruck**  
 Homepage , Lehrgang Personalbeurteilung , Ein Test zu Ihrer Menschenkenntnis

AKTIONSMENU: Vorheriger Nächste Inhalt Zurückverfolgen Neu laden Notizen machen Lesezeichen Suche Chat Foren

Bedienungsfeld Sichtbar für Design  
 Guided Tour  
 Designer-Plan  
 Neue Seite/Tool  
 Dateien verwalten  
 Kurs verwalten  
 Einstellungen  
 Inhaltsassistent

Kurs-Menü  
 Homepage  
 Lehrgang Personalbeurteilung  
 Kurs wiederaufnehmen  
 Diskussionsforen  
 Mailsystem  
 Chat  
 Visitenkarte Teilnehmen  
 Dateiaustausch  
 Whiteboard  
 Kalender

**Ein Test zu Ihrer Menschenkenntnis**

Dass es nicht einfach ist, Menschen nach dem äußeren Abbild zu beurteilen, bestätigt folgender Test aus der Zeitschrift FOCUS Ausgabe Nr.30/2000. Versuchen Sie an dieser Stelle herauszufinden, wer Künstler, wer Bürokrat, Menschenfreund, Mörder, Intelligenzbestie, Dummkopf, Pedant oder Schlampe, Tyrann oder Demokrat ist. Die Lösung wird sichtbar, wenn das Bild mit der Maus überfahren wird.




Abb. 7: Die Lernumgebung von WebCT (www.webct.com).

Derzeit befinden wir uns beim letzten Punkt der Evaluation - der Auswertung der Befragung der rund 450 Anwender(inne)n (Schüler(inne)n, Student(inn)en, Kursteilnehmer(innen)) von Testinstallationen der acht oben genannten Produkte an 20 verschiedenen Institutionen (Schulen, Fachhochschulen, Universitäten).

Das abschließende Ergebnis dieser letzten Evaluationsphase wird eine detaillierte Beschreibung der Anwendungserfordernisse dieser Lernplattformen sowie aufschlussreiche Nutzungserfahrungen der genannten Zielgruppen sein. Die Endergebnisse werden Ende Februar 2003 vorliegen.

## Literaturverzeichnis

Baumgartner, P/Häfele, H./Maier-Häfele, K. (2002a): Evaluierung von Lernmanagement-Systemen: Theorie - Durchführung - Ergebnisse. Handbuch E-Learning, Hohenstein, A./ Wilbers, K.(Hrsg.) : Fachverlag Deutscher Wirtschaftsdienst, Köln

Baumgartner, P/Häfele, H./Maier-Häfele, K. (2002b): E-Learning Praxishandbuch. Auswahl von Lernplattformen: Marktübersicht – Funktionen - Fachbegriffe. Innsbruck/Wien

Baumgartner, P./Frank, S. (2000): Der Mediendidaktische Hochschulpreis (MeDiDa-Prix) - Idee und Realisierung. In: Scheuerman, F.: Campus 2000 - Lernen in neuen Organisationsformen. Münster, S. 63-81

Baumgartner, P. Payr, S. (1997): Methods and practice of software evaluation: The case of the European Academic Software Award (EASA). In: Proceedings of ED-MEDIA 97 - World Conference on Educational Multimedia and Hypermedia. Charlottesville, AACE: S. 44-50.

Seeger, T. (1979): Die Delphi-Methode: Expertenbefragung zwischen Prognose u. Gruppenmeinungsbildungsprozessen. Freiburg

Thomé, D. (1988): Kriterien zur Bewertung von Lernsoftware. Heidelberg.

### **Webadressen:**

[W1] EASA: [www.easa.org/](http://www.easa.org/)

[W2] MeDiDa-Prix: [www.medidaprix.org/](http://www.medidaprix.org/)

[W3] Community Website: [www.virtual-learning.at/community](http://www.virtual-learning.at/community)