

# Expertenevaluierung durch die Methode des lauten Denkens am Beispiel eines Online-Lernspiels

Laura Ackermann, Melanie Heußner, Ludger Schmidt

Fachgebiet Mensch-Maschine-Systemtechnik

Universität Kassel

Mönchebergstr. 7

34125 Kassel

L.Ackermann@uni-kassel.de

M.Heussner@uni-kassel.de

L.Schmidt@uni-kassel.de

**Abstract:** Für den E-Learning-Anteil einer Weiterbildung wird ein browserbasiertes Lernspiel entwickelt, welches die Weiterbildungsteilnehmer im Selbststudium nutzen sollen. Die vorliegende Ausarbeitung stellt die Evaluation des Lernspiels durch ein interdisziplinäres Expertenteam dar. Durch die Methode des lauten Denkens konnten 108 Probleme des Lernspiels aufgedeckt werden.

## 1 Hintergrund und Problemstellung

Zur Entwicklung einer gemeinsamen Weiterbildungsmaßnahme für Pflegekräfte und Handwerker zum Thema Ambient Assisted Living (AAL) wird ein Blended-Learning-Ansatz verfolgt, der E-Learning-Phasen und Präsenzphasen kombiniert. Das Lernspiel basiert auf einem bestehenden ersten Entwurf [We13], dessen Inhalte für die geplante Weiterbildung neu erarbeitet wurden. Die durchgeführte Expertenevaluierung ist als erster Schritt der Qualitätssicherung einzuordnen.

Da die Ausbildungszeit der Zielgruppe unter Umständen mehrere Jahrzehnte zurückliegt und bisher keine oder kaum Erfahrungen mit computergestütztem Lernen vorhanden sind [He13], wurden bei der Evaluation sowohl Ergonomie-Kriterien als auch Aspekte der Didaktik, der Gestaltung und der Funktionalität überprüft.

## 2 Lösungsansatz

Die Entwicklung des Lernspiels erfolgte im Rahmen des menschenzentrierten Gestaltungsansatzes [DIN11]. Dabei handelt es sich um einen ganzheitlichen Gestaltungsansatz, der zur Steigerung der Gebrauchstauglichkeit eines Systems den gesamten Nutzungskontext und die Beziehungen und Wechselwirkungen seiner Komponenten betrachtet. Hierbei steht der Mensch im Mittelpunkt des Prozesses. Neben der Festlegung des Nutzungskon-

texts und der Nutzungsanforderungen gibt es mehrere Iterationen, in denen Gestaltungslösungen erarbeitet und evaluiert werden. Im Fall des Lernspiels wird die Evaluation in zwei Schritten durchgeführt: Im ersten Schritt, der hier dargestellt wird, lag der Fokus auf der Funktionalität, der Ergonomie, der Didaktik und der Gestaltung des Lernspiels. In einem zweiten, zukünftigen Schritt soll dann die Überprüfung der Praxistauglichkeit durch fachliche Experten aus dem Bereich AAL erfolgen, um inhaltliche Fehler aufdecken zu können. Dieses Vorgehen wurde gewählt, da technikbedingt Inhalte unproblematischer auszutauschen sind als Veränderungen am Design oder Verbesserungen der Funktionalität.

Als Evaluationsmethode wurde eine Kombination aus einer Expertenevaluation und der Methode des lauten Denkens gewählt. Eine expertenbasierte Evaluation zur Ermittlung von Usability-Problemen ist beispielsweise die heuristische Evaluation [NM90]. Diese wird zumindest in ihrer Basisversion nicht mit Nutzern durchgeführt [Le82, SB11]. Es lassen sich mehr Fehler aufdecken, wenn die Experten der Evaluation neben reinem Usability-Wissen zusätzlich auch über Fachwissen auf einem bestimmten Gebiet verfügen [Ni92]. Ein Nachteil der heuristischen Evaluation ist, dass nur Probleme erfasst werden, die durch bestehende Heuristiken abgedeckt werden. Bei der Entwicklung eines Systems ist es jedoch wichtig, auch scheinbar unbedeutende Probleme zu berücksichtigen, weil diese den Nutzer irritieren und den Gesamteindruck stören [Le82].

Die Methode des lauten Denkens wird häufig im Designprozess zur Evaluation von Prototypen angewendet [Ho05]. Sie ermöglicht die Gewinnung einer großen Anzahl wertvoller Hinweise auch mit wenigen Teilnehmern [SB11]. Durch Protokollierung verbalisierter Gedanken zu einer Aufgabe am entwickelten System wird es möglich, die Ergebnisse hinsichtlich der Systembedienung und der Eindrücke der Teilnehmer auszuwerten. Durch die Methode des lauten Denkens können somit Informationen über Gedanken und Emotionen während der Nutzung des Lernspiels gewonnen werden, die einen wichtigen Teil bei der menschenzentrierten Gestaltung darstellen.

Da der Fokus für diese Evaluation auf didaktischen, ergonomischen, funktionalen und gestalterischen Aspekten lag, wurde sie anstelle der sonst üblicherweise herangezogenen Nutzer [Le82, ES93, Ho05] mit Experten durchgeführt. Dabei wurden Experten befragt, die sowohl über Usability-Kenntnisse als auch über Wissen z.B. auf den Gebieten Design oder Informatik verfügen. Die Experten konnten in zwei Gruppen mit den beiden Schwerpunkten *Technik* und *Design* eingeteilt werden. Allen Experten war die Methode des lauten Denkens bereits bekannt. Durch Erfahrungen im Bereich der universitären Lehre kann bei allen Experten davon ausgegangen werden, dass didaktische Aspekte einfließen.

Als Orientierung zur Festlegung der Expertenzahl wurden sowohl Empfehlungen für heuristische Evaluationen (vier bis fünf Experten) als auch für die Methode des lauten Denkens (drei bis fünf Teilnehmer) herangezogen. Um beiden Empfehlungen gerecht zu werden und eine Tendenz über diese Empfehlungen hinaus erkennen zu können, wurde die Expertenzahl für diese Evaluation auf sechs festgelegt.

### 3 Durchführung und Auswertung

An zwei aufeinander folgenden Tagen wurde die Evaluation des Lernspiels mit sechs Experten in ihrer gewohnten Arbeitsumgebung durchgeführt. Sie verwendeten dazu die PCs, mit denen sie täglich arbeiten und somit vertraut sind. Während der Evaluierung wurden die Teilnehmer immer wieder durch den Evaluationsleiter ermutigt, ihre Gedanken und Eindrücke zu schildern. Die verbalen Äußerungen der Experten wurden wörtlich vom Evaluationsleiter protokolliert. Bei der Digitalisierung wurden alle Antworten, die sich auf ein Problem bezogen, in eine einheitliche Formulierung überführt.

Die weitere Auswertung erfolgte durch offenes Kodieren in Verbindung mit selektivem Kodieren [Ku07]. Die Problemnennungen wurden dabei den vier Kategorien *Ergonomie*, *Gestaltung*, *Funktionalität* und *Didaktik* zugeordnet. Hierzu nahmen zwei Auswertungspersonen zunächst getrennt voneinander eine Zuordnung aller Problemnennungen zu diesen Kategorien vor. Anschließend erfolgte ein Abgleich der Einordnungen, bei dem unterschiedliche Zuordnungen diskutiert wurden, so dass als Ergebnis eine eindeutige Kategorisierung der Problemnennungen erreicht wurde.

### 4 Ergebnisse

Insgesamt wurden durch die Evaluation mit den sechs Experten 108 verschiedene Probleme identifiziert. Abbildung 1 zeigt die Anzahl der Probleme für verschiedene Häufigkeiten der Problemnennungen. Dabei ist zu erkennen, dass 63 Probleme nur einmal erwähnt wurden und nur 4 Problemnennungen von allen Experten, also sechsmal, genannt wurden.

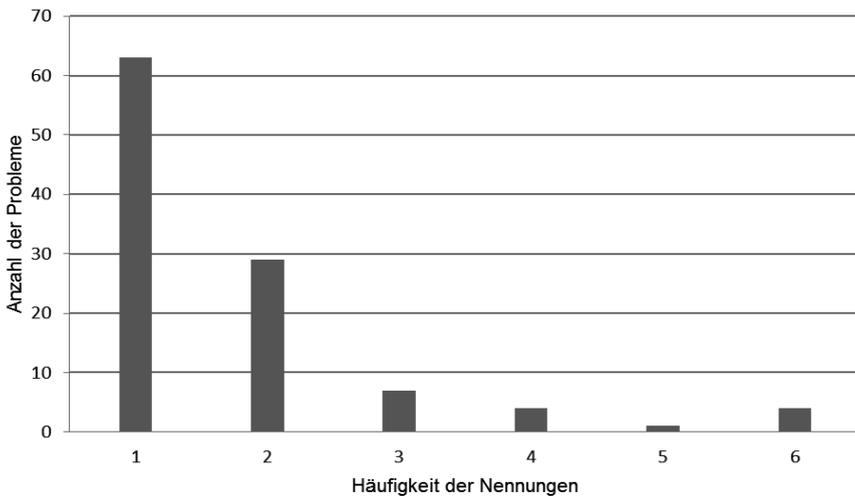


Abbildung 1: Häufigkeitsverteilung der Problemnennungen

Durch die Auswertung der aufgedeckten Probleme zeigte sich sowohl in der Kategorie *Gestaltung* als auch in der Kategorie *Ergonomie*, dass die meisten Probleme von der Gruppe der Experten im Bereich *Design* benannt wurden (Abb. 2). Auch die Zahl der zusätzlich von beiden Gruppen genannten Probleme liegt noch über den Nennungen durch die Expertengruppe *Technik*. Auffallend ist auch das Ergebnis in der Kategorie *Didaktik*; die Anzahl der Problemnennungen der Expertengruppe *Design* übersteigt die Nennungen der anderen Expertengruppe sowie die gemeinsamen Nennungen fast um den Faktor 3. Lediglich in der Kategorie *Funktionalität* sind die Werte beim Vergleich der beiden Expertengruppen ausgeglichen. Die insgesamt geringe Anzahl aufgedeckter Funktionalitätsprobleme erklärt sich daraus, dass es sich bereits um eine weiterentwickelte Version des Lernspiels handelt, in der mögliche Anfangsprobleme bereits behoben waren.

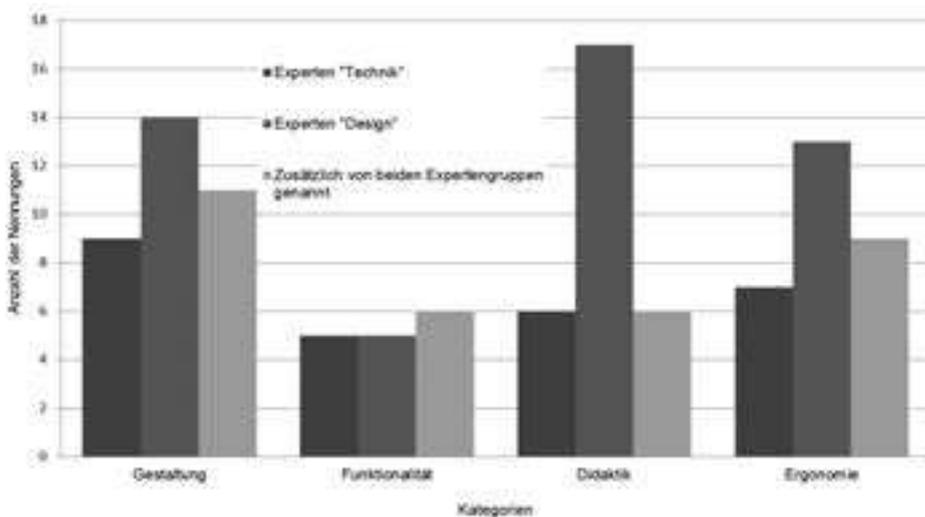


Abbildung 2: Problemnennungen der verschiedenen Expertengruppen

Es zeigt sich außerdem, dass die Anzahl der Problemnennungen pro Gruppe mit zunehmender Größe der Expertengruppe weniger stark zunimmt (Abbildung 3). Bei einer Gruppengröße von sechs Experten, wie sie in dieser Evaluation vorlag, lag die Anzahl der Problemnennungen bei 108. Eine Aussage darüber, welche Expertenanzahl am sinnvollsten für die Evaluation wäre, lässt sich aus den vorhandenen Daten jedoch nicht ableiten, da die Anzahl bereits im Voraus, basierend auf Empfehlungen für die heuristische Evaluation und die Methode des lauten Denkens, festgelegt worden war.

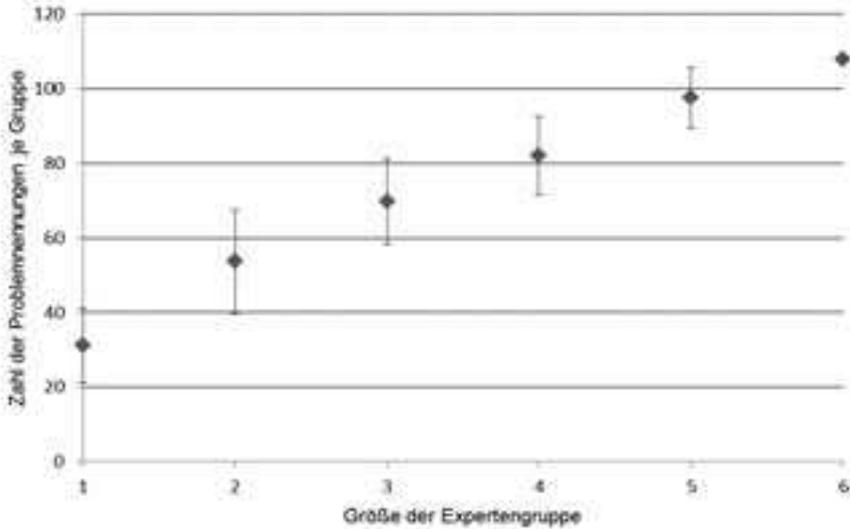


Abbildung 3: Problemnennungen bei unterschiedlicher Expertenanzahl

## 5 Diskussion und Fazit

Zur Evaluation eines Lernspiels wurde die Methode des lauten Denkens angewandt. Dabei wurden nicht – wie sonst üblich – Nutzer, sondern Experten aus den Bereichen *Technik* und *Design* befragt. Die Evaluation führte zu 108 ermittelten Problemen. Die meisten Probleme wurden einmalig genannt. Die festgelegte Anzahl von sechs Experten erwies sich für die Evaluation als geeignet und deckt sich mit den empfohlenen Zahlen für Nutzer bei der Methode des lauten Denkens bzw. für Experten bei der heuristischen Evaluation (vgl. Abschnitt 2). Die interdisziplinäre Zusammensetzung des Expertenteams, das sich in zwei Gruppen mit den fachlichen Schwerpunkten *Technik* und *Design* unterteilen ließ, erwies sich als sinnvoll. Obwohl die Expertise aller Teilnehmer auf den Gebieten der Usability und der Didaktik ähnlich war, konnten nicht nur die meisten Probleme in der Kategorie *Gestaltung*, sondern auch in den Kategorien *Ergonomie* und *Didaktik* durch die Expertengruppe aus dem Bereich *Design* aufgedeckt werden. Eine gewichtete Auswertung der Daten ist bisher in der angewendeten Methode nicht enthalten und stellt einen Nachteil gegenüber der heuristischen Evaluation dar. Eine Gewichtung der Probleme ist für weitere Evaluationen mit der Methode des lauten Denkens durch Experten zu empfehlen. Daneben ist zu überprüfen, inwiefern die Ergebnisse von einzelnen Personen, unabhängig von ihrer Expertengruppenzugehörigkeit, abhängen.

## 6 Ausblick

Um die Methode des lauten Denkens mit Experten überprüfen zu können, sind weitere Evaluationen geplant. Hierbei sollen die Problemnennungen der Experten gewichtet werden, um die Relevanz der Nennungen beurteilen zu können.

Im Gegensatz zu anderen Evaluationsmethoden, wie beispielsweise der heuristischen Evaluation, muss die Dokumentation nicht schriftlich erfolgen, sondern ist über ein Audio-Aufnahmegerät möglich. Somit entfällt der Aufwand, die eigenen Gedanken zu notieren, zumindest für den Evaluationsteilnehmer. Darüber hinaus können Evaluationen bei dieser Methode auch ohne persönliche Anwesenheit, z.B. über Chatfunktionen, durchgeführt werden. Durch den geringen Aufwand und die Flexibilität der Durchführung kann daher die Gewinnung geeigneter Experten für Evaluationen vereinfacht werden.

### Literaturverzeichnis

- [DIN11] DIN EN ISO 9241-210: Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme, 2011.
- [ES93] Ericsson, K. A.; Simon, H. A.: Protocol Analysis. Verbal reports as data. MIT Press, Cambridge, MA, 1993.
- [He13] Heußner, M. et al.: AAL-Weiterbildung für Pflege und Handwerk: Erste Ergebnisse einer Anforderungsanalyse. In: Verband der Elektrotechnik Elektronik Informationstechnik e. V. (Hrsg.): Lebensqualität im Wandel von Demografie und Technik: 6. Deutscher AAL-Kongress, Berlin 2013. VDE: Berlin, 2013, S. 286-290.
- [Ho05] Holzinger, A.: Usability Engineering Methods for Software Developers. In: Communications of the ACM - Interaction design and children, Volume 48 Issue 1. ACM: New York, 2005, S. 71-74.
- [Ku07] Kuckartz, U.: Einführung in die computergestützte Analyse qualitativer Daten. Verlag für Sozialwissenschaften: Wiesbaden, 2007.
- [Le82] Lewis, C.: Using the "Thinking Aloud" Method in Cognitive Interface Design. Technical Report RC-9265, IBM Thomas J. Watson Research Center: Yorktown Heights, 1982.
- [Ni92] Nielsen, J.: Finding usability problems through heuristic evaluation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York 1992. ACM: New York, 1992.
- [NM90] Nielsen, J.; Molich, R.: Heuristic Evaluation of User Interfaces. In: Proceedings of the CHI'90 Conference, Seattle 1990. ACM: New York, 1990. S. 249-256.
- [SB11] Sarodnick, F.; Brau, H.: Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendungen. Verlag Hans Huber, Bern, 2011.
- [We13] Wegener, R.: Smart SHK. [www.blendedcontent.de](http://www.blendedcontent.de), überprüft am 22.03.2013.