

Weißt Du es oder rätst Du noch? – ARS mit Rückmeldungen zur Antwortzuversicht bei Lernfragen

Prototypischer Vorschlag und erste Erkenntnisse

Lucas Braeschke¹, Iris Braun², Felix Kapp³ und Tenshi Hara⁴

Abstract: Die reine Erfassung von Antworten auf Lernfragen in Audience Response Systemen (ARS) leidet in der Regel unter einer Entkopplung zwischen tatsächlich vorhandenem Wissen, gefühltem Wissen und irrtümlich angenommenem Wissen. Für Lernende ist eine Einschätzung der Wissensqualität für erfolgreiches Lernen jedoch unverzichtbar. Durch die zusätzliche Erfassung der Antwortzuversicht wird Lernenden der tatsächliche Wissensstand zugänglich; darüber hinaus ist dann eine erfolgreiche Lernbedarfsermittlung möglich. Ob fehlendes oder falsches Verständnis vorliegt, oder Wissen nur partiell oder doch vollständig verarbeitet wurde, kann mit den Mitteln von Audience Response Systemen einfach und elegant verarbeitet und den Lernenden als individualisierte Auswertung zugänglich gemacht werden. In diesem Beitrag wird eine prototypische Implementierung mit ersten Evaluationsergebnissen präsentiert.

Keywords: Audience Response System, ARS, Zuversicht, Selbsteinschätzung

1 Einleitung

Audience Response Systeme (ARS) bieten die Möglichkeit, Lernende beim Wissenserwerb zu unterstützen. Durch Lernaufgaben, Umfragen und Abstimmungen lässt sich die Interaktivität in der Vorlesung erhöhen. Studierende haben die Möglichkeit, sich aktiv mit relevanten Konzepten auseinanderzusetzen und bei der Bearbeitung von Lernaufgaben auch Rückmeldung hinsichtlich ihres Lernfortschritts zu erhalten. Diese Information unterstützt sie beim erfolgreichen Regulieren ihres Lernprozesses: sie bekommen mit Hilfe von ARS Hinweise, welche Wissensgebiete bereits in der Lehrveranstaltung verstanden wurden und welche im Nachgang nochmals thematisiert werden sollten. Die häufig verwendeten Multiple-Choice-Fragen fordern die Lernenden dazu auf, die korrekte Antwort auf die Lernfrage auszuwählen, wobei das Potenzial dieser Lernaufgaben darin liegt, den Lernprozess sowohl in der aktiven Auseinandersetzung mit dem Inhalt während der Bearbeitung, als auch im sich anschließenden Reflexionsprozess, der durch die Beantwortung angestoßen werden kann, zu unterstützen [Ka11]. Die Lernenden

¹ Saxonia Systems AG, Fritz-Förster-Platz 2, 01069 Dresden, lucas.braeschke@saxsys.de

² Professur für Rechnernetze, Fakultät Informatik, TU Dresden, 01062 Dresden, iris.braun@tu-dresden.de

³ Institut für Psychologie und Arbeitswissenschaft, FG Mensch-Maschine-Systeme, TU Berlin, Fasanenstr. 1, 10623 Berlin, felix.kapp@tu-berlin.de

⁴ Berufsakademie Sachsen – Staatliche Studienakademie Dresden, Hans-Grundig-Str. 25, 01307 Dresden, tenshi.hara@ba-dresden.de

können die Rückmeldung verwenden, um ihren Lernprozess neu auszurichten, ihre Konzentration, Lernstrategien und Lernziele auf die aktuellen Gegebenheiten anzupassen. Wird neben der eigentlichen Multiple-Choice-Lernaufgabe noch zusätzlich die Antwortzuversicht abgefragt, kann dieser Vorgang effektiver gestaltet werden. Das Ziel ist es, besser auf Wissenslücken aufmerksam zu machen und somit auch den Lernprozess zu verbessern. So soll der Unterschied zwischen der eigenen Selbsteinschätzung und dem tatsächlich Gelernten aufgezeigt werden. Vor allem sind die Fälle, in denen diese beiden Faktoren divergieren, als besonders problematisch zu betrachten. Etwa wenn eine Frage trotz einer besonders hohen Zuversicht falsch beantwortet wurde oder aber wenn keine Zuversicht in die eigene richtige Antwort bestand. Diese Informationen sind ebenfalls für Lehrende wichtig. Nur wenn sie von etwaigen Problemen wissen, können sie diese auch adressieren. Vor diesem Hintergrund erscheint das Potenzial von Lernaufgaben ergänzende Zuversichtserfassung für die universitäre Lehre groß. Dabei können die erhobenen Daten sowohl unmittelbar in der Lehrveranstaltung zur Reflektion anregen und die Lehrenden unterstützen als auch zu einem späteren Zeitpunkt bei der Prüfungsvorbereitung und der Auswahl von Lernstoff und -strategien helfen.

Mit Auditorium Mobile Classroom Service (AMCS)⁵ werden kontinuierlich didaktische Konzepte mit den technischen Mitteln von ARS zusammengeführt. AMCS unterstützt dabei unterschiedliche Feedback-Mechanismen, u.a. auch Multiple-Choice-Fragen. Die Antworten auf die Fragen können als Datenbasis für die Lehrveranstaltungsevaluation, Lernstandanalysen sowie das Auslösen kognitiver und meta-kognitiver Prompts verwendet werden [Ka14]. Im Rahmen dieses Beitrags wird gezeigt, wie die zuvor genannte Zuversichtserfassung erfolgreich in ARS eingebracht und mit anschließender Zuversichtsbewertung Hinweise für die Prüfungsvorbereitung gegeben werden können.

2 Verwandte Arbeiten

Zuversicht (engl. *Confidence*) beschreibt laut Duden das „feste Vertrauen auf eine positive Entwicklung in der Zukunft“. Teilweise, wenn auch weniger verbreitet, wird in diesem Zusammenhang auch Sicherheit/Bestimmtheit (engl. *Certainty*) genutzt. Im Kontext der Lehre kann eine Frage nach der Zuversicht eingesetzt werden, um zu überprüfen, wie sicher sich die Teilnehmer bei der Beantwortung einer Frage sind.

Um die Zuversicht der Teilnehmer bei einer Lernfrage zu messen, wird meist eine vierstufige Skale genutzt (z.B.: „keine“, „geringe“, „mäßige“ und „hohe“ Zuversicht [Cu13],[LB16]). Die häufigste Anwendung findet die Zuversichtserfassung bei *Multiple-Choice-Fragen* (MCQ). Es existieren sowohl Varianten, bei denen die Sicherheit abgefragt wird, bevor die Antwortoptionen angezeigt werden (*Open Confidence*), als auch jene bei denen erst nach der Auswahl einer Antwort die Zuversicht erfasst wird (*Certainty-based Marking*).

⁵ <https://amcs.website/>

2.1 Zuversichtserfassung in Audience Response Systemen

In Audience Response Systemen wird eine Zuversichtserfassung bisher nur sehr selten umgesetzt. In einer Evaluation von 50 ARS und deren Funktionen [Ku19] wurde nur ein ARS identifiziert, das bisher Zuversichtserfassung umsetzt. Das an der TH Mittelhessen entwickelte *ARSnova* Projekt enthält in seiner vereinfachten Version *arsnova.click*⁶ auch die Abfrage der Zuversicht bei der Beantwortung von Lernaufgaben. Das System unterstützt fünf verschiedene Frage-Typen, wobei es sich bei vier um Lernfragen handelt. Die Zuversicht kann mittels einer Skala ausgewählt werden, welche in fünf Abstufungen unterteilt ist. Die Ausprägung auf der Skala ist neben der Beschriftung auch farbkodiert: *geraten* (dunkelrot), *sehr unsicher* (hellrot), *unsicher* (ocker), *relativ sicher* (hellgrün), *absolut sicher* (dunkelgrün). Die Antwortoption(en) und die Ausprägung der Zuversicht werden zeitgleich durch eine Schaltfläche abschließend bestätigt. Die Erfassung der Zuversicht ist sehr eingängig umgesetzt und lässt sich leicht bedienen und verstehen.

Wenn die Teilnehmenden die Frage beantwortet haben, wird für alle Nutzer eine Gruppenauswertung angezeigt, welche die gewählten Antworten und die durchschnittliche Zuversicht darstellt. Darüber hinaus beinhaltet die Gruppenauswertung ein Ranking der Teilnehmenden auf der Basis der Antwortzeiten. Dieses *Gamification*-Element konterkariert allerdings das didaktische Ziel der Zuversichtserfassung, nämlich Studierende dazu anzuregen, zu reflektieren, welchen Wissensstand sie in einem Wissensgebiet haben. Da schnelles Antworten im Ranking belohnt wird, ist eine tiefergehende Auseinandersetzung mit dem eigenen Wissenstand eher unwahrscheinlich.

2.2 Bestimmtheitsbasierte Markierung (Certainty-based Marking)

Certainty-Based Marking (CBM) ist in der Literatur auch unter dem Namen „*London Agreed Protocol for Teaching*“ bekannt. Entwickelt wird es seit 1994 am University College London (UCL). Lernende können dabei nicht nur mit einer Vielzahl an Testfragen Selbsttests durchführen, sie bekommen auch eine Rückmeldung inwiefern sie sich bereits präzise selbsteinschätzen können. In der webbasierten Umsetzung⁷ der UCL finden sich über 7000 Fragen. Den wichtigsten Teilbereich macht dabei der „*Medical and Biomedical Students’ Self-Test*“ mit über 5000 Einträgen aus.

Die Angabe der Zuversicht („*unsure*“, „*mid*“, „*sure*“, „*no idea*“ (Enthaltung)) erfolgt beim CBM zeitgleich zur Auswahl der Antworten auf die Lernfragen. Die Auswertung für die Lernenden beinhaltet eine Leistungskomponente (als *Präzision* bezeichnet; durchschnittlich erreichte Punkte) sowie eine Übersicht der Selbsteinschätzung (wie zuversichtlich die Studierenden bei der Beantwortung waren). Die Kombination beider Komponenten erlaubt eine Aussage über die Genauigkeit der Selbsteinschätzung der Lernenden zu machen. Dafür fließt die Zuversicht über den *CBM-Bonus* ein, welcher in Relation zum durchschnittlichen CBM-Wert für die entsprechende Präzision berechnet

⁶ <https://arsnova.click/>

⁷ <http://www.ucl.ac.uk/lapt/>

wird. Der Bonus wird vergeben, wenn der erzielte CBM-Wert nicht dem erwarteten durchschnittlichen Präzisionswert entspricht; er soll die abweichende Selbsteinschätzung in die Präzision einbeziehen. Dies gilt ebenfalls für den negativen Bereich, also wenn der eingeschätzte Wissensstand deutlich höher ist, als dies die Ergebnisse vorweisen. Der CBM-Bonus wird mit der Präzision zu einer allgemeinen Genauigkeit verrechnet. Die eigenen Ergebnisse im Kontext der Antworten anderer Nutzer zu sehen, bietet den Lernenden zusätzliche Möglichkeiten, sich selbst besser einzuschätzen.

2.3 Zuversichtsbasierte Markierung (Open Confidence-based Marking)

*The Open University*⁸ hat einen interessanten Ansatz im Kontext der Zuversichtseinschätzungen entwickelt, welcher online⁹ getestet werden kann. Die entwickelte Methode „*Open Confidence-based Marking*“ (OCBM) ist eine Variante des CBM. Der Unterschied beider Methoden liegt in dem Zeitpunkt der Einschätzung der Zuversicht. Ist bei CBM die Zuversicht erfasst worden, nachdem die Antwortoptionen eingesehen und ausgewählt wurden, wird dies bei OCBM durchgeführt, bevor diese angezeigt werden. Die Einschätzung der Zuversicht besteht wie bei der Umsetzung von UCL aus einer dreistufigen Skala („*low*“, „*medium*“, „*high*“), außerdem gibt es noch die Möglichkeit, mit „*give up*“ die Frage nicht zu beantworten. Zunächst sind nur der Fragetext und das Aufklappmenü für die Zuversicht zu sehen. Erst nachdem die Zuversicht ausgewählt wurde, werden die Antwortoptionen angezeigt. Nachdem eine Antwortoption ausgewählt wurde, wird das Feedback angezeigt. Darunter folgt eine Erläuterung zur richtigen Antwortoption. In der Auswertung wird für jede beantwortete Frage angezeigt, wie viele von den maximal zu erreichenden Punkten nach CBM-Schema erzielt wurden.

Durch die fehlenden Antwortmöglichkeiten zum Zeitpunkt der Selbsteinschätzung bezieht sich die Zuversichtserfassung auf ein größeres Wissensgebiet bzw. ist von generellerer Natur. Im Gegensatz zu Multiple-Choice-Aufgaben, bei denen man vorgegebene Antworten wiedererkennen oder ausschließen kann, beziehen sich die Einschätzungen hier auf Freitextaufgaben (da zum Zeitpunkt der Bewertung noch keine Antwortmöglichkeiten dargestellt werden). Es ist zu vermuten, dass die Aufgaben zu diesem Zeitpunkt als schwerer wahrgenommen werden und somit die Studierenden in ihrer Einschätzung auch unsicherer sind.

3 Erfassung und Bewertung der Zuversicht in AMCS

Wie in allen Audience Response Systemen bilden Live-Befragungen auch in AMCS die Kernfunktionalität [Ka14]. Jede Befragung besteht aus einer oder mehreren Lernaufgaben oder Umfragen, die die Studierenden über ihre eigenen Endgeräte (*Bring your own Device*) in Echtzeit beantworten können. Dabei wird zwischen folgenden Typen von Fra-

⁸ <http://www.open.ac.uk/>

⁹ <https://students.open.ac.uk/openma rk/t216-11b.demo.v03/>

gen unterschieden: Umfragen mit Einfachauswahl (*Single Choice*), Umfragen mit Mehrfachauswahl (*Multiple Choice*), Lernaufgaben mit Mehrfachauswahl und einer richtigen Antwort (*Single-Best Choice*) oder mehreren richtigen Antworten (*Multiple-Best Choice*), Freitext-Aufgaben (*Freetext*), Zuordnungsfragen (*Correct Assignment*), sowie Skalenfragen (*Scaled Questions*).

Im Falle von Lernaufgaben und Zuordnungsfragen erhalten die Studierenden direkt nach der Beantwortung neben der Information, ob die Frage richtig oder falsch beantwortet wurde, ein textuelles Feedback mit Hinweisen, warum einzelne Antworten richtig, nicht vollständig oder gar falsch sind. Im Unterschied zu anderen ARS bietet AMCS hier die zusätzliche Möglichkeit, die Frage ein weiteres Mal zu beantworten. Im Anschluss wird die richtige Lösung angezeigt, indem die falschen und die richtigen Antwortmöglichkeiten unterschiedlich farblich markiert werden. Dieser zweistufige Feedbackalgorithmus wurde mit dem Ziel implementiert, dass die Studierenden zur tieferen kognitiven Auseinandersetzung mit dem jeweiligen Inhalt angeregt werden.

Für die Integration der Zuversichtserfassung wurden diese Lernaufgaben ausgewählt, da hier eine direkte quantitative Auswertung der Ergebnisse und damit des Lernerfolges möglich sind. Somit können die Lernenden eine direkte Rückmeldung erhalten, ob und wieviel ihr Lernerfolg und ihre Selbsteinschätzung voneinander abweichen. Es wurden dabei sowohl Zuversichtserfassungen nach/während der Abgabe der Antworten (*Certainty Rating*) als auch vor der Anzeige der möglichen Antworten (*Open Confidence Rating*) prototypisch in AMCS umgesetzt. Die Auswertung der Aufgaben inkl. der notwendigen Zuversichtsbewertung erfolgte sowohl für einzelne Lernaufgaben als auch als aggregierte Werte über mehrere Aufgaben hinweg.

3.1 Lernfragen mit Zuversichtserfassung

Die Erfassung der Zuversicht während oder nach Abgabe der Antwort ermöglicht Lernenden eine genauere Einschätzung ihrer Leistung. Reicht das vorhandene Wissen aus, um die Frage mit hoher Sicherheit richtig zu beantworten, oder kann nur geraten werden? Die meisten der verwandten Arbeiten nutzen eine drei- oder fünfstufige Likert-Skala für die Eingabe der Zuversicht. In einer ersten Nutzerstudie wurden daher verschiedene Skalen für die Erfassung der Zuversicht in AMCS getestet und evaluiert, die Ergebnisse werden in Abschnitt 4 beschrieben. Schlussendlich wurde die Nutzung einer Prozent-Skala von 0 bis 100% ausgewählt, da dies eine genauere Eingabe der Zuversicht ermöglicht. Studierende können dazu einen Schieberegler verwenden, der nach der Frage und den möglichen Antworten eingeblendet wird (zur Umsetzung im AMCS-Prototyp siehe Abb. 1).

Bundestag

Es gehört nicht zu den Aufgaben des Deutschen Bundestages, ...

- Gesetze zu entwerfen.
- die Bundesregierung zu kontrollieren.
- den Bundeskanzler / die Bundeskanzlerin zu wählen.
- das Bundeskabinett zu bilden.

Wie sicher sind Sie sich, dass ihre Antwort richtig ist?

Zuversicht: 72%

Beantworten

Abb. 1: Multiple-Choice-Frage mit Zuversichtserfassung

3.2 Lernfragen mit offener Zuversichtserfassung

Im Gegensatz dazu wird beim *Open Confidence Rating* die Zuversicht bereits vor der Anzeige der möglichen Antworten abgefragt (siehe Abb. 2). Dies erfordert vom Lernenden eine etwas tiefergehende Auseinandersetzung mit der Fragestellung: Kennen sie die Antwort ganz sicher, weil sie das Themengebiet sehr gut verstanden haben, oder wissen sie nur eine ungefähre Antwort? Erst nach Auswahl der Zuversicht gelangen die Lernenden zur eigentlichen Ansicht der Fragestellung mit allen möglichen Antworten.

Bundespräsidentenwahl

Wer wählt in Deutschland den Bundespräsidenten / die Bundespräsidentin?

Wie sicher sind Sie sich, dass Sie diese Frage beantworten können?

Zuversicht: 50%

Weiter!

Abb. 2: Abfrage der Zuversicht vor der Beantwortung einer Lernfrage

3.3 Zuversichtsbewertung: Evaluation der Zuversichtserfassung

In der Literatur werden Lernende meist in vier Kategorien hinsichtlich der Kombination aus gezeigter Leistung und selbsteingeschätzter Zuversicht eingeteilt (etwa [Hu03], [Cu13]). Man kann dabei zwischen folgenden Kategorien unterscheiden:

Kein Verständnis (*uninformed*): In [Cu13] gilt ein Teilnehmer beim Beantworten einer Frage als „*uninformed*“, wenn dieser falsch geantwortet hat und angibt, sich unsicher zu sein. Dies zeigt nach [WG10] einen „*teaching moment*“ an, also eine Möglichkeit und einen Bedarf, mit Hilfe von Feedback und zusätzlichen Erläuterungen auf Lernende einzuwirken und so Wissenserwerb zu fördern.

Falsches Verständnis (*misinformed*): Der Status „*misinformed*“ beschreibt Lernende, welche die falsche Antwort ausgewählt haben, jedoch sehr zuversichtlich waren, die richtige gewählt zu haben [Cu13]. Dabei handelt es sich um eine fehlerhafte Selbsteinschätzung, die zu nicht adäquaten Strategien im Vorfeld von Prüfungen führen und das erfolgreiche Meistern dieser gefährden kann.

Partielles Verständnis (*partial knowledge*): Partielles Verständnis kann in zwei Fälle unterteilt werden [Bu02, Seite 807]. Grundsätzlich wurde die Frage richtig beantwortet, zugleich jedoch auch angegeben, sich mit der Antwort nicht sicher zu sein. Dies kann zum einen mangelndes Vertrauen in das eigene Wissen bedeuten. Zum anderen könnte es sich auch um den Versuch handeln, die richtige Antwort basierend auf dem lückenhaften Wissen zu erraten. Der erste Fall birgt die Gefahr, dass im Zuge der Prüfungsvorbereitung zu viel Zeit auf ein bereits gelerntes Thema verwendet wird.

Umfassendes Verständnis (*informed*): Dieser Fall beschreibt Lernende, welche die richtige Antwort auswählen und gleichzeitig angeben, sich mit der Antwort sicher zu sein. In diesem Fall kann davon ausgegangen werden, dass das entsprechende Wissen erlangt wurde und richtig angewendet werden kann.

Fragensammlung - Test - Lucas

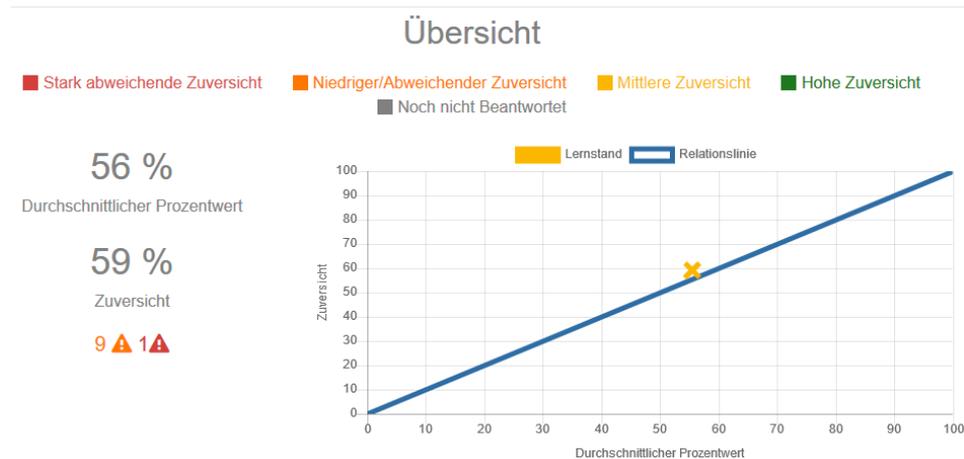


Abb. 3: Auswertung der Zuversicht (Ordinate) im Bezug zum Lernerfolg (Abszisse)

Die Auswertung in AMCS (siehe Abb. 3) hat zum Ziel, die Lernenden bei der Reflektion ihres Lernprozesses zu unterstützen und enthält die gleichen Informationen, welche auch

den vier vorgenannten Typen zu Grunde liegen. Durch eine farbliche Markierung angelehnt an die Ampelfarben [TT14] werden diese Abstufungen der Zuversichtseinschätzung verdeutlicht. Eine adäquate Selbsteinschätzung würde sich als linearer Fit von Zuversicht (Ordinate) und der durchschnittlichen Leistung in Prozent (Abszisse) darstellen (in Abb. 3 als blaue Linie dargestellt). Die eigene Zuversichtsbewertung gemittelt über alle beantworteten Fragen ist dann eine Koordinate mit Abweichung vom Fit, aus der die Lernenden leicht erkennen können, ob sie sich über- oder unterbewertet haben. Die Prozentwerte der Leistung unterscheiden sich für die verschiedenen Fragetypen in ihrer Ausprägung. Für MBC- und CA-Fragen können Werte zwischen 0 und 100% erreicht werden, bei SBC nur 0 oder 100 %. Je mehr Fragen unterschiedlicher Art beantwortet wurden, umso repräsentativer sind die errechneten Werte.

Der beispielhaft dargestellte Lernstand in Abb. 3 ist somit sehr nah am Ideal für leicht über das mittlere Niveau erreichte korrekte Antworten. Die Zuversicht zeigt eine minimale Tendenz zur Selbstüberschätzung. Die Koordinate ist deshalb so nah am Fit, weil beim mittleren Answererfolg eine dazu passende mittlere Zuversicht vorliegt (in Zahlen: 56% Erfolgsquote bei 58% Zuversicht).

3.4 Fragenkatalog

Basierend auf der Zuversichtsbewertung wird ein Fragenkatalog sowohl für die gezielte Wiederholung von Problemfragen (keine oder nicht vollständig richtige Antworten) als auch für die Korrektur von fehlerhafter Wissensschätzung (hohe Zuversicht bei falscher Antwort oder niedrige Zuversicht bei richtiger Antwort) erstellt. Idealerweise wird der Fragenkatalog den Lernenden zur Auswahl gestellt, wodurch diese selbst Schwerpunkte festlegen können. Um die Lernenden bei der Auswahl der zu wiederholenden Fragen zu unterstützen, empfiehlt sich eine Markierung der Wiederholungsdringlichkeit. Im AMCS-Prototyp wird entsprechender Handlungsbedarf farblich markiert (siehe Abb. 4). Auch hier wurden in Anlehnung an die Ampel-Metapher Rot und Orange als Signalfarben für erhöhten Handlungsbedarf und Gelb und Grün für weniger Dringlichkeit ausgewählt. Zusätzlich wird die Anzahl der Fragen in der jeweiligen Kategorie mit einem Ausrufezeichen hervorgehoben. Grau markierte Fragen wurden noch nicht beantwortet.

4 Nutzerstudien

Um eine möglichst intuitive, den Anforderungen der Lernenden gerechte Umsetzung der Zuversichtserfassung zu erreichen, wurden unterschiedliche, auf Visualisierung und Nutzbarkeit ausgelegte Ansätze untersucht. Dazu wurden Probandenbefragungen mit Papier-Prototypen durchgeführt, auf Basis derer Ergebnisse erste prototypische Änderungen am AMCS-Prototyp vollzogen wurden. Diese wurden dann auf Basis des User Experience Questionnaire¹⁰ (UEQ) vergleichend gegenübergestellt. Zusätzlich wurden

¹⁰ <https://www.ueq-online.org>

die Probanden bezüglich Farbgestaltung und Symbolauswahl befragt. Insbesondere wurde erfasst, ob und wie sich der Arbeitsfluss verbessern ließ.

Wiederholung



Lernangebote



Lernstand

19

Anzahl der Fragen



Start

Wählen Sie die Fragen für die Wiederholung aus.

■ Stark abweichende Zuversicht
 ■ Niedrige/Abweichende Zuversicht/Prozentwert
 ■ Mittlere Zuversicht und Prozentwert
■ Hohe Zuversicht und Prozentwert
 ■ Noch nicht Beantwortet

✓ ✗ **Organe der Bundesrepublik Deutschland**

53% Durchschnittlicher Prozentwert	86% Zuversicht	32% Zuversichts Abweichung	1 ▲ 1 ▲
<p>Gesetzgebung</p> <p>Die deutschen Bundesländer wirken an der Gesetzgebung des Bundes mit durch ...</p> <p>(Einfachauswahl mit einer richtigen Antwort)</p> <p>100% Richtig 92% Zuversicht</p>	<p>Bundestag</p> <p>Es gehört nicht zu den Aufgaben des Deutschen Bundestages, ...</p> <p>(Einfachauswahl mit einer richtigen Antwort)</p> <p>0% Richtig 100% Zuversicht</p>		
<p>Gewaltenteilung</p> <p>Ordnen sie die Organe den Gewalten zu!</p> <p>(Zuordnungsfrage)</p> <p>33% Richtig 50% Zuversicht</p>	<p>Exekutive</p> <p>Was gehört in Deutschland zur Exekutive?</p> <p>(Mehrfachauswahl mit mehreren richtigen Antworten)</p> <p>80% Richtig 100% Zuversicht</p>		

Abb. 4: Farblich markierte Vorschläge für Wiederholungsaufgaben entsprechend der gegebenen Antworten und der zugehörigen Zuversichtserklärungen (grün: gute Einschätzung ... rot: kritisch)

In einer ersten Studie mit fünf Probanden wurde erfasst, wie die Zuversicht bei der Erfassung visualisiert werden kann. Es wurden drei Papierprototypen vorgeschlagen: Doktorhüte (umso mehr, umso zuversichtlicher), Skala (0 bis 100%; höhere Prozentwerte entsprechen höherer Zuversicht) und Emojis (der Zuversicht entsprechende Gesichtsausdrücke vom Frowny bis zum Smiley).

Anschließend wurden die UEQ-Werte für die drei Prototypen erfasst (siehe Tab. 1), mit der Option weitere Probanden im Falle großer Abweichungen zu befragen. Die Doktorhüte erreichen dabei auf allen Dimensionen schlechtere Werte im Vergleich zu den anderen beiden Visualisierungen, welche wahlweise präferiert werden. Basierend auf den zusätzlich erfassten freien Rückmeldungen der Probanden kann die fehlende Originalität in Kombination mit schlechter Identifikation mit den Doktorhüten als negativer Faktor identifiziert werden.

Den Probanden wurden drei Skalenversionen präsentiert: in 1%-, 10%- und 20%-Schritten. Diesen einfachen Skalen mangelt es eindeutig an Ästhetik. Im Bereich der Neuheit wurden sie sogar negativ bewertet. Dies widerspricht gängigen Erfahrungen, da Skalen

sonst ein viel verwendetes und daher bekanntes Visualisierungsinstrument in ARS sind und ist eventuell auf die kleine Stichprobengröße zurückzuführen. Die Emojis haben eine hohe Attraktivität, Verlässlichkeit und Neuheit sowie zeichnen sich durch eine deutlich geringere Varianz aus, als die anderen Prototypen. Dies macht sie offensichtlich zur geeignetsten Visualisierung – zumindest für AMCS.

	Doktorhüte		Skala		Emojis	
	AVG	VAR	AVG	VAR	AVG	VAR
Attraktivität	1,3	1,41	0,83	2,33	1,93	0,43
Klarheit	1,4	1,64	1,95	0,39	1,7	1,23
Effizienz	1,85	0,68	2	0,59	1,35	1,27
Verlässlichkeit	1,6	1,21	1,8	0,73	1,3	0,61
Stimulation	1,4	1,52	0,65	1,58	1,7	0,42
Neuheit	0,85	1,02	-0,05	4,08	1,8	0,98

Tab. 1: Auswertung zu den drei Prototypen
(Skalen von -3 – „entsetzlich schlecht“ bis +3 – „extrem gut“; entsprechend UEQ¹⁰)

Nach der Integration eines Emoji-Schiebereglers (siehe Abb. 1 und Abb. 2) wurde eine zweite Befragung durchgeführt. Ziel war es, Probleme bei der Eingabe und dem Arbeitsfluss der Fragenbeantwortung und Zuversichtserfassung zu identifizieren. Dazu wurde der Prototyp der AMCS App in einer echten Lehrveranstaltung eingesetzt und im Anschluss ein UEQ¹⁰ von den 14 Studierenden ausgefüllt. Mit einer durchschnittlichen Klarheit von 2,3¹¹ ist ein deutliches Indiz für die Eignung dieser Visualisierungsform gegeben. Basierend auf freien Rückmeldungen der Studierenden kann festgestellt werden, dass die Emojis selbsterklärend und einfach zu bedienen sind. Die von uns implementierte Kombination aus Emojis mit Schieberegler¹² erlaubt darüber hinaus eine feingranulare Zuversichtserfassung gegenüber der alleinigen Verwendung von Emojis.

Für die dritte Studie wurde die Erstellung einer Fragensammlung auf Grundlage der Leistungsdaten und Zuversichtsbewertung im Prototyp implementiert. Damit sollte untersucht werden, ob und wie die Vorschläge den Studierenden bei der Prüfungsvorbereitung helfen. Zehn Probanden wurden gebeten sich in eine fiktive Lehrveranstaltung zu versetzen, in welcher sie 20 Fragen zu Politik, Geografie und Geschichte Deutschlands (entnommen aus dem *Einbürgerungstest*) bearbeiteten. Im Anschluss bekamen alle Probanden die Möglichkeit, sich eine Übersicht ihrer Ergebnisse anzuschauen und Fragensammlungen nochmals in Vorbereitung auf eine zukünftige Prüfung zu bearbeiten. Für diese Wiederholungsfragen standen unterschiedliche, vom Prototyp automatisch generierte Fragenkataloge zur Auswahl: 1. alle Lernfragen, unabhängig vom Antworterfolg und/oder Zuversicht, 2. keine Wiederholung, 3. Fragen mit abweichenden Bewertungen (insbesondere für *misinformed* und *partial knowledge*), 4. nicht richtig beantwortete

¹¹ Die Tabelle zur UEQ-Auswertung fehlt hier aus Platzgründen.

¹² Auf Vorschlag der Probanden aus der ersten Befragung; durch Verschieben des Reglers ändert sich das Emoji in der Schubschaltfläche.

Fragen, 5. Fragen mit Antworten niedriger Zuversicht, 6. Fragen mit Antworten hoher Zuversicht, sowie 7. noch nicht beantwortete Fragen.

In Tab. 2 sind exemplarisch die Ergebnisse der Probanden aufgelistet. Die Tabelle zeigt neben dem Erfolg bei der Bearbeitung der Aufgaben und der erfassten Zuversicht auch die von den Probanden wahrgenommene Eignung der Darstellung der Ergebnisse. Die Frage zielte darauf ab, ob sich die Probanden von der Darstellung des Lernstandes repräsentiert fühlten (0 - „überhaupt nicht“, 7 - „vollkommen“). Die fünfte Spalte zeigt, welchen Fragenkatalog die Probanden für die Wiederholung verwendet haben. Die empfundene Nützlichkeit der Fragenkataloge wurde ebenfalls bewertet (0 – „ungeeignet“ bis 7 – „absolut geeignet“).

Proband	Antwort- erfolg	Zuversicht	Reprä- sentation	verwendeter Fragenkatalog	Nützlichkeit
1	45%	55%	6	1	7
2	69%	65%	6	4	7
3	57%	53%	5	1	6
4	50%	69%	6	1	7
5	65%	74%	6	4	6
6	77%	84%	6	4; 5	k. A.
7	64%	51%	3	3; 5; 6	6
8	84%	54%	6	5	7
9	74%	62%	5	4	5

Tab. 2: Ergebnisse der Evaluation des Fragenkataloges zur Prüfungsvorbereitung

Die Probanden wählten für die Prüfungsvorbereitung zwischen den drei Strategien: 1) alle Lernfragen noch einmal wiederholen, 4) alle nicht korrekt beantworteten Fragen, oder 5) alle Fragen mit einer geringen Zuversicht noch einmal bearbeiten. Das Minimieren von Wiederholungen (Katalog 2) und die Auswahl noch nicht bearbeiteter Fragen (Katalog 7) wurden von keiner Person gewählt. In den drei häufigsten Strategien findet sich Fragenkatalog 5, welcher als Grundlage die Zuversichtsbewertung für die Auswahl von Wiederholungsfragen nutzt. Aufgrund der geringen Probandenzahl können die Ergebnisse aber nicht als repräsentativ gewertet werden.

5 Zusammenfassung und Ausblick

Durch die Zuversichtserfassung ist es wesentlich leichter möglich, den Wissensstand der Teilnehmenden beim Beantworten von Lernfragen einzuschätzen. Dies kann in verschiedenen Kontexten wie dem Selbstregulierten Lernen (SRL) hilfreich sein, um die Lernstrategie anzupassen. Die zusätzliche Angabe der Zuversicht hinsichtlich der korrekten Bearbeitung der Aufgabe fördert die Reflexion des eigenen Wissensstandes und regt zu einer tieferen Auseinandersetzung mit der Lernfrage an. Darüber hinaus kann mit der Zuversichtsbewertung im Anschluss aufgezeigt werden, ob die Beurteilung des eige-

nen Wissenstandes korrekt war. Inwieweit die Zuversichtserfassung eine messbare Verbesserung des Lernerfolges oder eine Anpassung der Lernstrategie der Studierenden zur Folge hat, soll in weiteren Experimenten mit größeren Stichproben untersucht werden.

Literaturverzeichnis

- [Br18] Braun, I. et al: Technology-enhanced self-regulated learning: Assessment support through an evaluation centre. Proc. IEEE 42nd Annual COMPSAC, Symposia CELT: Computer Education & Learning Technologies, Tokyo. 2018.
- [Bu02] Burton, R. F.: Misinformation, partial knowledge and guessing in true/false tests. *Medical Education*, 36(9):805–811. 2002.
- [Cu13] Curtis, D. A. et al: Does student confidence on multiple-choice question assessments provide useful information? *Medical Education*, 47(6):578–584. 2013.
- [Ga06] Gardner-Medwin, A.: Confidence-Based Marking - towards deeper learning and better exams, Seiten 141–149. Routledge. 2006.
- [Hu03] Hunt, D. P.: The concept of knowledge and how to measure it. *Journal of Intellectual Capital*, 4(1):100–113. 2003.
- [Ka14] Kapp, F. et al: Meta-cognitive Support in University Lectures Provided via Mobile Devices - How to Help Students to Regulate Their Learning Process during a 90-minute Class. Proc. 6th CSEdu, Seiten 194–199. ScitePress. 2014.
- [Ka11] Kapp, F.; Körndle, H.: Was lerne ich aus einer Lernaufgabe? a) gar nichts, b) Faktenwissen, c) etwas über meine Lernstrategien, d) Antwort b und c sind richtig. In T. Köhler & J. Neumann (Hrsg.), *Wissensgemeinschaften. Digitale Medien - Öffnung und Offenheit in Forschung und Lehre* (S. 178-187). Münster: Waxmann.
- [Ku19] Kubica, T. et al: Choosing the appropriate audience response system in different use-cases. Proc. 10th ICETI, Orlando. 2019.
- [LB16] Luetsch, K.; Burrows, J.: Certainty rating in pre- and post-tests of study modules in an online clinical pharmacy course - A pilot study to evaluate teaching and learning. In: *BMC Medical Education* 16.1, Seite 267. 2016.
- [TT14] Tak, S.; Toet, A.: Color and Uncertainty: It is not always Black and White. In Elmqvist, N., Hlawitschka, M., Kennedy, J., editors, *EuroVis - Short Papers*. The Eurographics Association. 2014.
- [WG10] Wakabayashi, T.; Guskin, K.: The Effect of an “Unsure” Option on Early Childhood Professionals’ Pre- and Post-Training Knowledge Assessments. In: *American Journal of Evaluation* 31.4, Seiten 486–498. 2010.
- [Zi00] Zimmerman, B. J. et al: Attaining self-regulation: a social cognitive perspective. *Handbook of self-regulation*, 13. 2000.