

Mixed Reality Konzepte für Online Gemeinschaften

Mixed-reality concepts for online communities

Welche Gamification motiviert? Ein Experiment zu Abzeichen, Feedback, Fortschrittsanzeige und Story

*Athanasios Mazarakis, Paula Bräuer
Christian Albrechts Universität zu Kiel, Institut für Informatik*

Zusammenfassung

Gamification soll durch Spielelemente die Motivation in spielfremden Kontexten erhöhen, an Untersuchungen zur Wirkung einzelner Spielelemente mangelt es jedoch. Die vorliegende Studie leistet einen Beitrag dazu diese Lücke zu schließen. In einem Experiment mit 505 Teilnehmern konnte gezeigt werden, dass die Spielelemente Abzeichen, Feedback, Fortschrittsanzeige und Story einzeln und kombiniert angewendet zu signifikanten Motivationssteigerungen führen, sich aber im Motivationspotenzial unterscheiden.

1 Einleitung und verwandte Arbeiten

Gamification ist ein Konzept, welches zur Motivationssteigerung für die Erfüllung bestimmter Aufgaben angewendet wird [1]. Dieses Konzept ist nach Deterding u. a. [2] als die Verwendung von Spielelementen in einem spielfremden Kontext definiert. Unter einem spielfremden Kontext können zum Beispiel Systeme zur Förderung der Fitness oder des Schulunterrichts fallen [3]. Zu den am meisten untersuchten Spielelementen zählen unter anderem Punkte, Bestenlisten und Level [1, 4, 5]. In der vorliegenden experimentellen Feldstudie werden die vier Spielelemente Abzeichen, Feedback, Fortschrittsanzeige und Story isoliert untersucht. Die verwendeten Spielelemente sollen nun kurz betrachtet werden.

Abzeichen sind virtuelle Artefakte, welche visuell dargestellt werden. Sie werden an den Nutzer für die Erledigung von Aufgaben verliehen [6]. Hamari [7] konnte in einer zweijährigen Studie nachweisen, dass Anwender einer gamifizierten Variante einer Sharing-Plattform deutlich aktiver waren als ohne Gamification.

Feedback ist eines der wichtigsten Spielelemente für den Bereich Gamification [4]. Feedback soll dem Nutzer Informationen über seine Leistung oder den Status der Anwendung liefern, wodurch eine Änderung des Verhaltens möglich ist [8]. Hierbei wurde schon in mehreren Onlinestudien gezeigt, dass selbst einfachstes Feedback motivierend wirken kann [9].

Die akute Forschungslücke zum Spielelement Fortschrittsanzeige macht eine zuverlässige Einschätzung des Potenzials eher schwierig. So konnte Sailer [3] zeigen, dass nur 2 % der verwendeten Spielelemente eine Fortschrittsdarstellung sind, wohingegen Abzeichen, Punkte und Bestenlisten am häufigsten vorkommen. Auch bei Dicheva u. a. [10] zeigt sich, dass Fortschrittsanzeigen in Relation zu anderen Spielelementen ein Nischendasein fristen.

Schließlich ist auch das Story-Element für Gamification essenziell, besonders wenn es darum geht, etwas zu lernen oder Anweisungen zu erteilen, da es Dingen Bedeutung verleiht [8]. Nach Sailer u. a. [3] kann durch eine Geschichte einerseits Interesse geweckt werden, welches dann auf einen anderen Kontext übertragen werden kann. Andererseits kann durch diverse Entscheidungsmöglichkeiten in der Geschichte, ein motivierend wirkendes Autonomieempfinden entstehen. Zusätzlich haben Geschichten das Potenzial positive Gefühle auszulösen [11].

Eine Einschränkung, die bei vielen der experimentellen Studien im Bereich der Gamification auftritt, besteht darin, dass keine Einzelbetrachtung der Elemente vorgenommen wird, sondern es werden alle gleichzeitig angewendet, was zu einer Konfundierung führt [5]. Die vorliegende Studie will diesen Zustand beheben.

2 Rahmenbedingungen des Experiments

Im Rahmen der vorliegenden Studie wurde ein Experiment durchgeführt, welches eine systematische und erweiterte Version einer anderen Studie [12] ist. Diese untersuchte die Spielelemente Abzeichen und Story. Allerdings wurden in der Studie zusätzlich zu den untersuchten Spielelementen das Element Feedback kombiniert. So wurde möglicherweise die Wirkung der getesteten Spielelemente verfälscht. Einen weiteren Mangel stellt die Gestaltung des Spielelements Story dar, welches sehr schlicht gehalten wurde und auch nicht zu statistisch signifikanten Ergebnissen führte. Mit der vorliegenden Studie sollen diese Schwachpunkte behoben werden.

Das Ziel beider Studien besteht darin, zu untersuchen, ob sich die Anzahl an beantworteten Fragen in einem Quiz durch den Einsatz von verschiedenen Spielelementen steigern lässt. Grundsätzlich ging es im Quiz der anderen Studie [12]

darum, Fragen zu Kontinenten und Ländern zu beantworten. Im Experiment der vorliegenden Studie wurde der Fragenkatalog der Ursprungsstudie größtenteils übernommen und um 20 Fragen zum Thema Weltraum erweitert. Diese Fragen wurden zur Änderung des Elementes Story verwendet. Zusätzlich wurde als Spielelement eine Fortschrittsanzeige hinzugefügt.

Das durchgeführte Experiment umfasste insgesamt sieben Versuchsbedingungen: Neben einer Kontrollgruppe (KG) ohne weitere Manipulation werden sechs Experimentalgruppen mit den folgenden Spielelementen und Kombinationen unterschieden: Feedback (FB), Fortschrittsanzeige (PB), Fortschrittsanzeige und Feedback (PB+FB), Story (ST), Story und Feedback (ST+FB) sowie Abzeichen und Feedback (BA+FB). Die Kombinationen aus dem Spielelement Feedback und den anderen Spielelementen wurde konträr zu [5] gewählt, um die Vergleichbarkeit dieser Studie und der Ursprungsstudie [12] zu gewährleisten. Aufbauend auf diese Rahmenbedingungen wurden folgende sechs Hypothesen aufgestellt:

- Hypothese 1 (H1): Die Gruppe in der Experimentalbedingung Feedback beantwortet mehr Fragen als die Kontrollgruppe ($FB > KG$)
- Hypothese 2 (H2): In der Experimentalbedingung Fortschrittsanzeige werden mehr Fragen beantwortet als die Kontrollgruppe ($PB > KG$)
- Hypothese 3 (H3): Die Gruppe in der Experimentalbedingung Story beantwortet mehr Fragen als die Kontrollgruppe ($ST > KG$)
- Hypothese 4 (H4): Die Kombination Fortschrittsanzeige und Feedback beantwortet mehr Fragen als nur Feedback ($PB+FB > FB$)
- Hypothese 5 (H5): Die Kombination Story und Feedback beantwortet mehr Fragen als nur Feedback ($ST+FB > FB$)
- Hypothese 6 (H6): Die Kombination Abzeichen und Feedback beantwortet mehr Fragen als nur Feedback ($BA+FB > FB$)

2.1 Beschreibung des Quiz

Das Quiz besteht aus 190 Fragen mit jeweils vier Antwortmöglichkeiten. Die Fragen des Quiz sind in neun Blöcke thematisch unterteilt. Sieben Blöcke befassen sich mit je zehn allgemeinen Fragen zu den sieben Kontinenten und jeweils fünf Fragen zu vier Ländern auf dem Kontinent, ausgenommen von Australien und Antarktika. Die zwei anderen Blöcke beinhalten je zehn Fragen zu Mond und Weltraum. Die Fragen werden in allen Versuchsbedingungen in derselben Reihenfolge angezeigt. Ein Beispiel für eine Frage zeigt Abbildung 1.



Abbildung 1: Beispielfrage aus der Experimentalgruppe Story

Die Probanden werden unter keiner der Versuchsbedingungen über den Umfang des Quiz in Kenntnis gesetzt. Die Grundidee des Quiz ist es, die Probanden so lange spielen zu lassen, wie es ihnen Spaß bringt, um anhand der Anzahl beantworteter Fragen die Wirkung der verschiedenen Spielelemente zu messen.

Für die Umsetzung des Feedback Elementes wird ein „richtig-falsch“ Feedback gewählt. Die vom Probanden gewählte Antwort färbt sich direkt nach Beantwortung einer Frage grün oder rot. Zusätzlich wird bei einer falschen Antwort die richtige Lösung grün markiert. Die Abzeichen sind identisch zur Ursprungstudie [12]. Abbildung 2 stellt die erreichbaren Abzeichen dar.



Abbildung 2: Darstellung aller erreichbaren Abzeichen

Acht der 14 Abzeichen werden nur bei korrekter Beantwortung einer bestimmten Frage vergeben. Die restlichen Abzeichen werden bei einer hohen Antwortgeschwindigkeit von richtigen Antworten oder für mehrere korrekte Antworten in Folge freigeschaltet. Für beide Kategorien werden jeweils drei Abstufungen verwendet, und zwar von leicht zu schwer (zum Beispiel drei, fünf oder zehn richtige Fragen hintereinander). Diese werden zuerst nur grau hinterlegt. Es ist nicht ersichtlich, welche Bedingungen man erfüllen muss, um ein Abzeichen zu erhalten. Dadurch soll nur das Verhalten belohnt werden, das Quiz zu spielen und nicht das Bemühen, ein Ziel zu erreichen. Nachdem ein Abzeichen erspielt wurde, konnte man durch die Ansteuerung mit dem Mauszeiger erfahren, wofür das Abzeichen vergeben wird. Zusätzlich wurde man durch eine Bildschirmeinblendung informiert, dass man gerade ein Abzeichen erhalten hat.

Für das Element Story wird eine Geschichte mit einem Außerirdischen gewählt. Dieser Außerirdische wurde von einer Weltraumkommission auf die Erde gesandt, um zu prüfen, ob der Planet einer Weltraumautobahn weichen soll. Der Proband soll nun Fragen beantworten, um den Außerirdischen vom Wert der Erde zu überzeugen. Zu allen in den Fragen thematisierten Kontinenten und Ländern wurden passende Bilder erstellt, die die Story visuell unterstützen sollen. Beim Wechsel zum nächsten Kontinent bzw. Land wird ein neues Bild und ein Text, mit einem einleitenden Kommentar des Außerirdischen angezeigt. Als Zwischenereignis wurde ein Fragenblock zum Thema Mond eingebaut. Der Außerirdische teilt dem Spieler mit, dass der Mond gerade gesprengt werden soll. Um dies zu verhindern, müssen weitere Fragen beantwortet werden. Das Zwischenereignis soll die Geschichte abwechslungsreicher gestalten und die Motivation des Spielers konstant halten. Der Spieler erhält keine Information über den Ausgang der Geschichte. Dadurch soll vermieden werden, dass Probanden noch einmal am Quiz teilnehmen, um den Ausgang der Geschichte zu ändern.

Für die Fortschrittsanzeige wird aufgrund der sehr themenbezogenen Fragen, ein rundes Design in Form eines sich füllenden Erdballs gewählt. Das Bild der Erde wird in 20 Schritten stückweise aufgedeckt. Der Proband sieht über der Fortschrittsanzeige, wie viele Frage er noch zu beantworten hat, bis das nächste Stück des Erdballs aufgedeckt wird. Die Anzahl der Fragen, die zu beantworten sind, ist nicht linear angestiegen, sondern variiert zwischen 3 und 15 Fragen. Insgesamt wurden die 190 Fragen auf 20 Stücke der Fortschrittsanzeige verteilt.

3 Die Feldstudie

Das Experiment wurde im Zeitraum von März und April 2017 über 33 Tage durchgeführt. Die Probanden wurden online über soziale Netzwerke wie Facebook und Xing akquiriert. Für die Teilnahme wurde keine Entlohnung gezahlt und auch nicht mit einer solchen geworben. Final nahmen 531 Probanden an der Studie teil. Hiervon wurden 20 aufgrund von doppelt auftretenden IPs und Session-IDs entfernt. Weitere sechs Einträge wurden wegen widersprüchlicher Angaben zu Alter und Geschlecht, sowie durch zusätzliches auffälliges Verhalten, wie z. B. zu schnellem Durchklicken, von der Auswertung ausgeschlossen. Dadurch verblieben insgesamt 505 Probanden für die statistische Auswertung.

Die Zuweisung der Probanden zu den Versuchsbedingungen wurde randomisiert vorgenommen. Ein Wechsel der Gruppe durch den Probanden war nicht möglich. Zu Beginn des Quiz wurden die Teilnehmer über die Anonymisierung der Daten sowie über die Tatsache aufgeklärt, dass zum Absenden der Daten ein „Quiz beenden“-Button gedrückt werden musste. Auch darauf, dass nur solange Fragen zu beantworten waren „wie man will“, wurde hingewiesen. Anschließend wurden freiwillige Angaben zu Alter und Geschlecht abgefragt.

Die Angaben über Alter und Geschlecht ergaben eine Beteiligung von 61 % Frauen und 34 % Männern, 5 % machten keine Angaben. Das durchschnittliche Alter lag bei 31.47 Jahren. Der Median lag bei 27 Jahren und die Altersangaben reichten von 14 Jahren bis 78 Jahren. Die Unterschiede im Alter zwischen den Frauen und Männern sind nicht statistisch signifikant. Umgesetzt wurde das Experiment mit PHP und einem MySQL-Server.

Wurde das Quiz durch drücken des „Quiz beenden“-Button oder durch beantworten aller 190 Fragen abgeschlossen, erhielt der Proband die Information, wie viele der Fragen er insgesamt richtig beantwortet hatte. Außerdem wurde die Möglichkeit eines Kommentars über ein Kommentarfeld gegeben.

4 Ergebnisse

Es werden nun zuerst die deskriptiven Ergebnisse und danach die inferenzstatistischen Ergebnisse dargestellt. Die Verteilung der Probanden auf die Versuchsbedingungen ist in Abbildung 3 dargestellt.

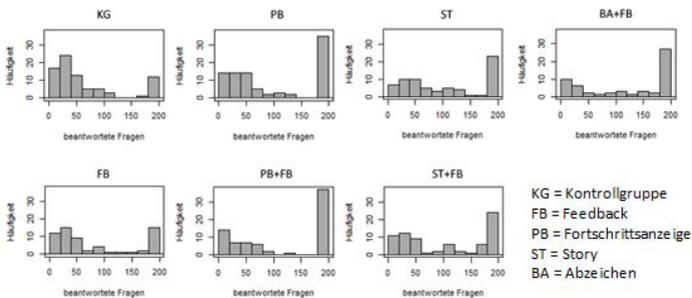


Abbildung 3: Häufigkeitsverteilung beantworteter Fragen je Bedingung

In allen Gruppen liegt eine bimodale Verteilung vor. Auffällig sind insbesondere die Versuchsbedingungen mit der Fortschrittsanzeige (PB, PB+FB) und den Abzeichen (BA+FB). So haben exakt 50.0 % der Probanden in der Versuchsbedingung mit der Fortschrittsanzeige und Feedback, bzw. 39.3 % in der mit der Fortschrittsanzeige, alle Fragen beantwortet. In der Bedingung Abzeichen und Feedback waren es auch bemerkenswerte 41.4 %. Unabhängig von den inferenzstatistischen Ergebnissen kann bereits festgestellt werden, dass die Fortschrittsanzeige, bzw. die Abzeichen, jeweils immer in Kombination mit Feedback, die Probanden am besten dazu motivierten, alle Fragen zu beantworten.

Im Schnitt wurden 98.79 Fragen beantwortet, mit einer Standardabweichung von 74.29. In Tabelle 1 sind sowohl die Anzahl der Probanden, Mittelwert der beantworteten Fragen, zugehörige Standardabweichung (SD) und die Anzahl der Probanden, welche alle Fragen beantworteten je Bedingung angegeben. Außerdem unterscheiden sich die einzelnen Versuchsbedingungen nicht statistisch signifikant in Bezug auf das Verhältnis von der Anzahl der beantworteten Fragen und der Anzahl der korrekt beantworteten Fragen mit $F(6, 498) = 1.73, p = .113$. Das Erfolgserlebnis korrekt beantworteter Fragen hat in keiner Versuchsbedingung zu signifikant mehr oder weniger beantworteten Fragen geführt, womit auch eine Konfundierung mit den Versuchsbedingungen nahezu ausgeschlossen werden kann.

Bereits aus den deskriptiven Ergebnissen ist zu erkennen, dass die Teilnehmer der Kontrollgruppe im Durchschnitt weniger Fragen beantworteten als in den anderen Versuchsbedingungen. Für eine genauere und belastbarere Interpretation der Ergebnisse folgt nun die inferenzstatistische Analyse.

Für die statische Auswertung der Ergebnisse wird eine Varianzanalyse verwendet, um die Mittelwertvergleiche zwischen den unterschiedlichen Gruppen interpretieren zu können. Der Test zur Homogenität der Varianzen (Levene-Test) für die Anzahl der beantworteten Fragen wird statistisch signifikant mit $p = .000$, wobei die Levene-Statistik 7.37 beträgt. Bei allen folgenden Ergebnissen wird daher von nicht gleichen Varianzen ausgegangen und entsprechend konservativ korrigiert. Dadurch kommt es auch zu krummen Werten bei den Freiheitsgraden.

Tabelle 1: Anzahl, Mittelwert, Standardabweichung (SD) und die Anzahl der Probanden mit allen gespielte Fragen je Bedingung

Bedingung	Anzahl	Mittelwert	SD	Alle Fragen beantwortet
KG	80	63.35	61.58	12
FB	62	82.61	71.07	24
PB	89	101.21	76.04	35
PB+FB	74	114.78	78.44	37
ST	69	105.59	69.16	23
ST+FB	74	106.39	73.39	24
BA+FB	57	123.44	75.54	25

Die Varianzanalyse zeigt einen statistisch signifikanten Unterschied zwischen den einzelnen Versuchsbedingungen an, $F(6, 498) = 5.68, p = .000$. Da die Homogenität der Varianzen nicht gegeben ist, muss durch den Welch-Test entsprechend korrigiert

werden [13]. Nach der Korrektur lautet das Ergebnis der Varianzanalyse $6.51, p = .000$. Durch den statistisch signifikanten Unterschied in der Anzahl der beantworteten Fragen zwischen den einzelnen Versuchsbedingungen, können alle aufgestellten Hypothesen nun statistisch untersucht werden und einseitig getestet werden. Der Vergleich der Mittelwerte der Feedback-Gruppe mit denen der Kontrollgruppe liefert ein signifikantes Ergebnis, $t(121.01) = 1.70, p = .046$. Folglich kann H1 gestützt werden und davon ausgegangen werden, dass das Feedback Element die Probanden motiviert. Auch der Vergleich der Mittelwerte der Gruppe mit Fortschrittsanzeige mit den Mittelwerten der Kontrollgruppe ergibt ein signifikantes Ergebnis, $t(165.26) = 3.57, p = .000$. Damit kann davon ausgegangen werden, dass die Fortschrittsanzeige eine motivierende Wirkung hat, H2 wird gestützt. Die Probanden in der Bedingung mit der Story haben ebenfalls statistisch signifikant mehr Fragen beantwortet als die der Kontrollbedingung, $t(137.47) = 3.91, p = .000$. Damit kann auch H3 gestützt werden.

Alle drei Bedingungen, die neben dem Feedback Element ein weiteres Element verwendeten, liefern beim Vergleich der Mittelwerte, mit denen der Gruppe, die nur das Feedbackelement benutzte, signifikante Ergebnisse: PB+FB $t(133.15) = 2.51, p = .007$; ST+FB $t(131.19) = 1.92, p = .029$; BA+FB $t(114.58) = 3.03, p = .002$. Folglich können auch H4, H5 und H6 gestützt werden, in Kombination wirken also die Elemente besser als nur Feedback alleine.

5 Fazit, Diskussion und Ausblick

Die vorliegende Studie hat zu neuen Ergebnissen im Bereich der Gamification Forschung geführt. In einem Experiment mit 505 Probanden konnte gezeigt werden, dass die Spielelemente Fortschrittsanzeige, Abzeichen und Story auch einzeln angewendet eine signifikant motivierende Wirkung aufweisen. Auch der angenommene Effekt, der durch die Verwendung eines zusätzlichen Elementes erzeugt werden sollte, konnte nachgewiesen werden.

Verglichen mit den Ergebnissen der Ursprungsstudie konnte die Aussage über die motivierende Wirkung der Abzeichen bestätigt werden. Die Annahme, dass das Story Element alleine angewendet nicht motiviert, konnte jedoch widerlegt werden. Dies ist wahrscheinlich auf die Umgestaltung der Story zurückzuführen, was den Schluss zulässt, dass die Wirkung des Elementes stark von der Umsetzung abhängt und somit nicht verallgemeinerbar ist. Insbesondere der Einsatz von Fortschrittsanzeigen und Abzeichen kann aufgrund der sehr positiven Ergebnisse zur Steigerung der Motivation empfohlen werden.

Eine Limitation der Studie liegt beim Design der Fortschrittsanzeige vor. Eine Bedingung, die in allen Versuchsgruppen gegeben sein sollte, bestand darin, den Probanden keine Informationen über die Länge des Quiz bereitzustellen, damit diese nicht versuchen, das Maximum an Fragen zu beantworten. Durch die Verwendung der Fortschrittsanzeige war jedoch eine grobe Schätzung über die Gesamtzahl der Fragen möglich.

Die Wirkung der Spielelemente ist trotz allem stark davon abhängig, in welchem Kontext bzw. System sie eingesetzt werden. Genauso hat auch die Umsetzung der Elemente einen Einfluss auf deren Effekt, wie man am Beispiel der Story sehen kann. Während noch in der ersten Studie die Umsetzung nicht ideal und wenig motivierend umgesetzt wurde [12], ist dies in der vorliegenden Studie nun besser gelöst worden. Diese Einschränkung bezüglich der Umsetzung existiert aber in allen wissenschaftlichen Studien. Durch die Anwendung der *ceteris paribus* Klausel wurde versucht, dieser Einschränkung entgegenzuwirken. Eine Generalisierung der Ergebnisse der Feldstudie wird jedoch nicht angenommen.

Literaturverzeichnis

- [1] Seaborn, K., Fels, D.I., *Gamification in Theory and Action: A Survey*. International Journal of Human-Computer Studies 2015, 74, 14–31.
- [2] Deterding, S., Dixon, D., Khaled, R., Nacke, L., *From Game Design Elements to Gamefulness: Defining Gamification*, in: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments - MindTrek '11, ACM, New York, New York, USA 2011, pp. 9–15.
- [3] Sailer, M., *Die Wirkung von Gamification auf Motivation und Leistung: Empirische Studien im Kontext manueller Arbeitsprozesse*, Springer, Wiesbaden 2016.
- [4] Zichermann, G., Cunningham, C., *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*, O'Reilly Media, Sebastopol, California, USA 2011.
- [5] Mekler, E.D., Bruehlmann, F., Tuch, A.N., Opwis, K., *Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance*. Computers in Human Behavior 2017, 71, 525–534.
- [6] Antin, J., Churchill, E., *Badges in Social Media: A Social Psychological Perspective*, in: CHI 2011 Gamification Workshop Proceedings, New York, New York, USA 2011.
- [7] Hamari, J., *Do Badges Increase User Activity? A Field Experiment on the Effects of Gamification*. Computers in Human Behavior 2017, 71, 469–478.

-
- [8] Kapp, K.M., *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education*, Pfeiffer, San Francisco, California, USA 2012.
- [9] Mazarakis, A., Using Gamification for Technology Enhanced Learning: The Case of Feedback Mechanisms. *Bulletin of the IEEE Technical Committee on Learning Technology* 2015, 4, 6–9.
- [10] Dicheva, D., Dichev, C., Agre, G., Angelova, G., Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society* 2015, 18, 75–88.
- [11] Sailer, M., Hense, J., Mandl, H., Klevers, M., Psychological Perspectives on Motivation Through Gamification. *Interaction Design and Architecture(S)* 2013, 19, 28–37.
- [12] Mazarakis, A., Gamification: Eine experimentelle Untersuchung der Spielelemente Abzeichen und Story. *Mensch & Computer 2017* in Druck.
- [13] Field, A., *Discovering Statistics Using SPSS*, SAGE, London 2013.