

Valenzen in Serious Games

Spielerisch auf neuen Wegen der UX-Messung



Insa Wulf
Hochschule der Medien Stuttgart
insa.wulf@gmx.net

Dr. Huberta Kritzenberger
Hochschule der Medien Stuttgart
kritzenberger@hdm-stuttgart.de

Abstract

Serious Games sind Videospiele, in denen Spieler nicht ausschließlich spielen, sondern gleichzeitig lernen. Warum ist es hier wichtig, die Ux messen zu können? User Experience kann darauf hindeuten, ob die Verknüpfung von Spiel- und Lerninhalt so erfolgreich ist, dass bei den Spielern konzentrationsfördernder Flow entsteht. Ein Zugang zur Bewertung von UX sind von den Nutzern erlebte Wertigkeiten (Valenzen) von UX-Merkmalen. In einer explorativen Studie wurden Valenzen erhoben, um die verschiedenen UX-Merkmale von Serious Games betrachten zu können. In dieser Studie setzten die Spieler selbst die Valenzen bei verschiedenartigen Merkmalen: von Grafik über Humor zu Schwierigkeitsgrad sowie Sound und mehr. Diese Valenzen wurden retrospektiv besprochen, indexiert und nach Häufigkeiten gewichtet ausgewertet. Der Einsatz von Valenzen für die UX-Messung in Serious Games ermöglicht es, diejenigen Merkmale zu dokumentieren, welche für die Spieler in Bezug auf ihre Erwartungen relevant waren und so die Entwicklung des Spielflows bestimmt haben.

Keywords:

/// Valenzen
/// Serious Games
/// User Experience Research
/// User Experience

Einleitung

Wäre es nicht angenehm, in der Zukunft nicht nur Bücher oder eBooks zu lesen, sondern die multimedialen Möglichkeiten von Videospiele zu nutzen, um zu lernen? Dieser Artikel stellt eine Studie vor, dessen Ziel es war, das Nutzungserlebnis eines sogenannten „Serious Games“ zu validieren. Serious Games bieten sich als Medium an, in dem am Computer gespielt und gleichzeitig gelernt werden kann. Zu welchem Thema die Weiterbildung stattfindet, das hängt natürlich vom jeweiligen Spiel ab. Dieser Artikel befasst sich nun damit, wie das Nutzungserlebnis, also die UX, bei einem so lebendigen Produkt wie einem Computerspiel gemessen werden kann. In einer explorativen Studie wurde eine Messung anhand von Valenzen und einer Indizierung ebendieser durchgeführt. Der genaue Ablauf der Studie wird in den folgenden Absätzen dargestellt. Verglichen wurde diese empirisch erhobene UX zusätzlich mit der Vorerfahrung der Spieler und einer zuvor durchgeführten Umfrage zu dem Genre, welches in dem Serious Game bedient wird.

Warum sollte das Nutzungserleben bei Serious Games gemessen werden?

Serious Games sind wie beschrieben Videospiele, in denen Spieler nicht ausschließlich spielen, sondern sich gleichzeitig weiterbilden. Die Spiele vermitteln die zu lernenden Inhalte beispielsweise über die Hintergrundgeschichte und ihre Handlungsabläufe, wodurch die Grenzen zwischen Lern- und Spielinhalten verschwimmen. In den Momenten, in denen im Spiel die Lernsituation für die Spielenden nicht offensichtlich, sondern primär das Spielen wahrgenommen wird, entsteht im Spielflow starke Konzentration (Fritz, 1997, S.212). Aufgrund dieser angenehmen Möglichkeit, spielerisch konzentriert zu lernen, erhält diese Art von Videospiele derzeit zunehmend Aufmerksamkeit.

Welchen Mehrwert bietet an dieser Stelle nun die Messung von User Experience bei der Entwicklung von Serious Games? Kann positive User Experience nachgewiesen werden, ist das ein Hinweis darauf, dass die Verknüpfung von Spiel- und Lerninhalt so gelungen ist, dass die Lernsituationen

unbewusst bleiben und bei den Spielenden konzentrationsfördernder Flow entsteht (Fritz, 1997). Wird negative User Experience gemessen, ist der Flow folglich ein schwer zu erreichender Zustand. Ist es gegenteilig doch möglich geworden, dass Spieler das Flow-Erlebnis erlangen, entsteht ein Entwicklungs- oder Lernprozess (Csíkszentmihály, 1992, S.70f). Eine hohe gemessene UX kann also auf entstehenden Spielflow und damit auf einen erleichterten Lernprozess in dem Serious Game hindeuten. Diese Hinweise bilden für die Entwicklung und für die kontinuierliche Verbesserung eines Serious Games eine fundierte Grundlage für Optimierungsansätze.

Wie sieht der Ansatz der explorativen Studie aus?

Ein möglicher Zugang um die UX zu bewerten, sind Wertigkeiten (Valenzen) von UX-Merkmalen. Diese Wertigkeiten des Nutzungserlebens bei dem Benutzen eines Produkts resultieren aus einem Gefühl, welches sich einstellt, sobald Frustration oder Erfüllung von Grundbedürfnissen erlebt wird. Dieses rangiert auf

der Gefühlsdimension „positives Gefühl“ bis „negatives Gefühl“, der sogenannten Valenzdimension (Hassenzahl, 2008). Diese Valenzen erleben auch die Spielenden bei Serious Games und werden daher gemessen um die Wahrnehmung der UX zu ermitteln und dadurch auf den Spielflow schließen zu können.

Wie verlief die Studie?

In der explorativen Studie wurde den 9 Teilnehmern kein Leitfaden mit speziellen Handlungsanweisungen und Aufgaben vorgegeben, wie es zum Beispiel in einem Usability Test üblich ist. Stattdessen durften die Spielenden das Serious Game frei explorieren. Die Studie wurde so angelegt, da die Unterhaltung, das Flow-Erleben, alltäglichem Computerspielen meistens als einzige Motivation zugrunde liegt. Freies Computerspielen kann nicht in einem linearen Handlungsablauf vorhergesagt werden, da jedes Explorieren des Spielraums verschieden verläuft. Diese Freiheit ist Teil des Spielerlebens.

Nach einer Einleitung wurde den Teilnehmern von einem Moderator der Valenzmarker vorgestellt, mit dessen Hilfe die Teilnehmer bei einem positiv oder negativ erlebten Merkmal Wertigkeiten anzeigen können, dieser ist in Abbildung 1 dargestellt. Bei den Merkmalen kann es sich um verschiedenartige Spielelemente wie Grafik, Schwierigkeitsgrad, Humor und vieles mehr handeln.

Im Anschluss spielten die 9 Teilnehmer je frei für 20 Minuten das Lernadventure „Winterfest“. Der Schauplatz dieses Computerspiels, welches den Serious Game Award in Gold im Jahr 2011 auf der CeBIT gewonnen hat, ist eine Stadt im Mittelalter. Von dort muss der Protagonist mit seinem Kumpanen, einer Ratte, den Weg zurück in unsere Gegenwart finden und viele Rätsel lösen. Das Spiel soll Erwachsenen die Möglichkeit bieten, Lesen, Schreiben und Rechnen zu üben. Hierfür kann „Winterfest“ auf der Website <http://www.lernspiel-winterfest.de/> heruntergeladen und kostenlos gespielt werden. Während der Spielphase von „Winterfest“ setzten

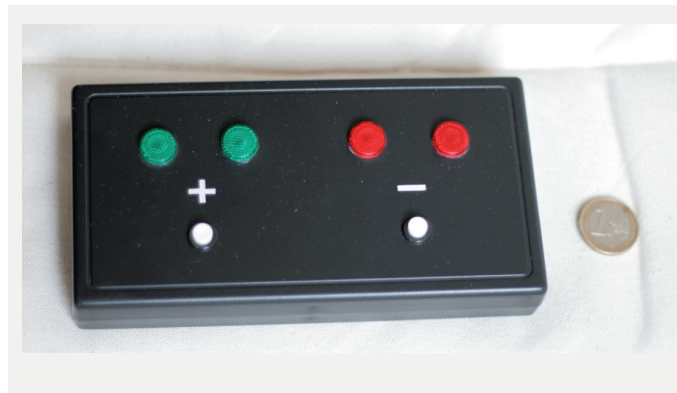


Abb. 1.
Valenzmarker zum
Markieren positiver (+)
oder negativer (-)
Valenzen

die Spieler selbständig die Valenzen bei verschiedenartigen Merkmalen mithilfe des Valenzmarkers. Während dieser Spielphase wurden der Bildschirm mit dem Computerspiel inklusive der Blickbewegungen des Spielenden mittels Eyetracker sowie der Valenzmarker mit einer weiteren Videoaufnahme aufgenommen. [Abb. 1]

In einer anschließenden retrospektiven Befragungsphase betrachteten der Teilnehmer und der Moderator gemeinsam das Video der Spielaufzeichnung mit den Blickbewegungen und die Aufnahme des Valenzmarkers (siehe Abbildung 2). Hatte der Teilnehmer während des Spielens einen Marker gesetzt, wurde das Video gestoppt um die Spielsituation zu besprechen. Diese Videonachbesprechung orientiert sich an dem Retrospective Thinking Aloud (De Jong, Van den Haag, 2002), im Folgenden mit RTA abgekürzt, und der Valenzmethode (Burmester, Jäger, Festl, Mast, 2011). Bei der RTA geben die Teilnehmer retrospektiv wieder, was sie zum jeweils im Video wiedergegebenen Zeitpunkt gedacht haben.

Eine alternative Möglichkeit hierzu wäre das Concurrent Thinking Aloud (De Jong, Van den Haag, 2002), im Folgenden CTA genannt, bei dem die Spielenden parallel zur Handlung ihre Gedanken erläutern. Das CTA wurde nicht angewandt, da es eventuell den natürlichen Spielprozess gestört hätte, welches in einer Studie zu erkennen ist, in der De Jong und Van den Haag (2002) RTA und CTA verglichen haben. Dabei haben die Teilnehmer, welche CTA anwandten, weniger erfolgreich Testaufgaben

bewältigt als die Teilnehmer, welche retrospektiv verbalisierten. De Jong und van den Haag weisen darauf, dass dieser Unterschied in dem Arbeitsaufwand liegen könne, den das beim CTA erforderliche, laute Denken verursacht und sich damit negativ auf das Leistungsvermögen auswirken könne. Während der retrospektiven Befragungsphase des Serious Games wurden außerdem ausdrucksstarkes Verhalten und Mimik besprochen, wie zum Beispiel Lachen oder spontane Ausrufe der Spielenden, welche der Moderator während der Spielphase notiert. Diese Stellen wurden bei der Videonachbesprechung zusätzlich zu den Markern thematisiert, um herauszufinden, welche Elemente des Spiels diese Reaktionen hervorgerufen haben.

Es fällt auf, dass ein Teilnehmer während der Spielphase wenig Marker setzte. Obwohl er bei einer vorhergehenden Übung normal viele Marker gesetzt hat, wird er nicht während der Spielphase dazu aufgefordert, mehr Marker zu setzen. Dies geschieht deshalb nicht, um ihn nicht im Spielfluss zu stören. Denn während des Computerspielens kann der Effekt eintreten, dass die Spielenden im Flow vollkommen ihre Aktivität und ihr Bewusstsein verschmelzen und so ihre komplette Aufmerksamkeit auf das Spiel lenken, so dass sie sich nicht mehr selbst reflektieren sowie nicht mehr über andere Dinge außerhalb der Spielwelt nachdenken (Fritz, 1997). Tritt dieser Effekt bei dem Serious Game ein und vergisst ein Teilnehmer dadurch Marker zu setzen, sollte der Moderator die Teilnehmer in dem möglichen Fall nicht durch eine Erinnerung

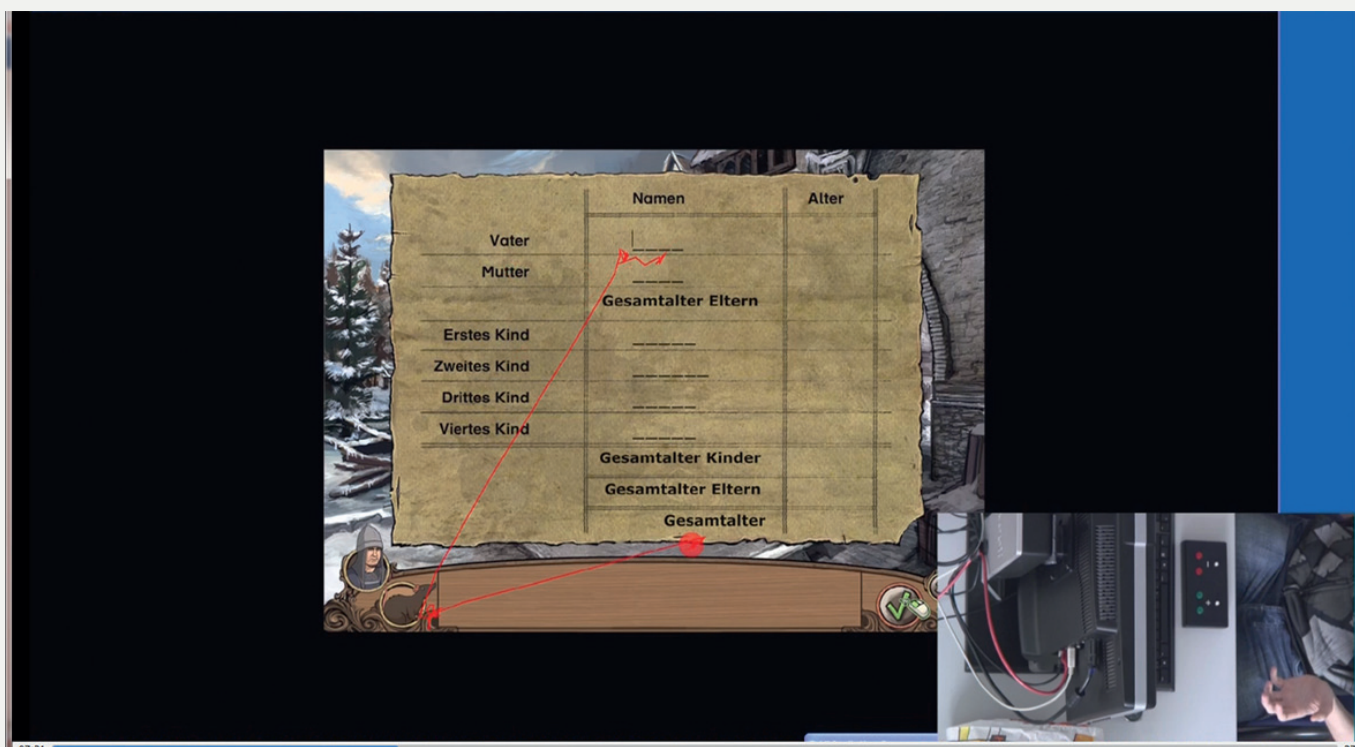


Abb. 2.
Aufgezeichnete Minispiel-Szene mit Gaze Plots und Aufnahme des Valenzmarkers

an das Markersetzen aus diesem Zustand lösen. Stattdessen wird vermehrt das Verhalten, Mimik und Ausrufe beobachtet, sowie während der RTA vertieft Aufmerksamkeit darauf gelenkt, wie der Teilnehmer bestimmte neue Schauplätze, Rätsel oder Dialoge empfand. So kann auch ohne Marker noch hinreichend herausgefunden werden, wie das Erleben der Teilnehmer in der Spielphase war. [Abb. 2]

In der RTA wurde auf die Befragungstechnik des „Ladderings“, welche der Valenzmethode zugrunde liegt (Burmerster, Jäger, Festl, Mast, 2011), verzichtet. Diese Fragetechnik wurde dahingehend entwickelt, die unbewusst hinter einer positiv oder negativ empfundenen Eigenschaft stehenden Grundbedürfnisse zu ermitteln. In der Studie stand im Vordergrund, die positiven oder negativen UX-Merkmale zu dokumentieren und mit den Erwartungen

abzugleichen. Daher besprachen Moderatorin und Teilnehmer in der RTA die Eigenschaften der Elemente des Serious Games, welche veranlasst hatten, den Valenzmarker zu drücken. Wenn es die Spielsituation zulaß, verglichen sie diese Spielelemente außerdem mit anderen bereits gespielten Computerspielen, aus denen die Spielenden ihre Erfahrungswerte gebildet hatten.

Wie wurde ausgewertet?

Die Ergebnisse der retrospektiven Befragung wurden in Tabellen protokolliert. Diese wurden segmentiert, orientierend an den von den Teilnehmern gesetzten Markern. Um darzustellen, welche Segmente ähnliche Merkmale thematisieren, wurden diese testpersonenübergreifend indexiert und anschließend nach Themen und Spielsituation sortiert. Eine spielübergreifende Zusammenfassung schloss

dies ab, wie beispielhaft in Abbildung 3 in einem Ausschnitt der Tabelle dargestellt wird. Die Index-Nummern generieren sich aus der Teilnehmer-ID und der jeweiligen Listen-Nummer des gesetzten Markers. Dies ermöglichte eine schnelle Identifizierung der jeweiligen Teilnehmer sowie eine grobe Einordnung der Indizes in den Ablauf des Serious Games. Bei einer Gesamtzahl von durchschnittlich 45 Indizes pro Teilnehmer wird empfohlen im Anschluss anhand der quantitativen Daten auffallende Indizes zu wählen und primär diese qualitativ nach positiver und negativer Valenz auszuwerten. [Abb. 3]

Was galt als UX Merkmal?

Die Teilnehmer setzten positive sowie negative Valenzen bei sehr verschiedenen Spielelementen. Diese Spielelemente konnten derart tiefe Wirkung bei den

Grafik Buchansicht schön	Details und allgemein immer bei Büchern schön							0810
Grafik coole Hintergrundgrafik	sieht cool aus; verlassene Atmosphäre; erinnert an anderes gern gespieltes Spiel Gene Machine; generell wichtig (09), erinnert an Monkey Island; schöne Design und Atmosphäre	0301 und 0303 und 0315	0446 und 0450	0610 und 0623	0711 und 0713 und 0719 und 0720 und 0721	0804 und 0809	0903 und 0915 und 0925	
Grafik der Karte	liebevoll und mittelalterlich; blinkt bei neuen Orten; ungewohnt, besser von oben direkt Standort sehen; gut erkennbar wo hingehen möglich; inkonsistent Kreuz unten rechts; Standort nicht erkennbar	0160	0426 und 0448	0520	0622	0826 und 0717 0828	0914	
Grafik des Hauptmenüs	schon und professionell							0802
Grafik Dialogansicht	Eindruck, dass Person 2 Mal im Raum entsteht, weil selbe Darstellung, verwundernd; wichtiges Gespräch mit Ratte nicht erkennbar, Rücken zu Betrachter gedreht; süße Ansicht bei den Dialogen			0514		0830 und 0840	0931	
Grafik Differenz zwischen Audio und Grafik	am Marktplatz viele Geräusche und Gemurmel, aber nichts zu sehen; Geräusche, Kommentare und Grafik passen nicht zusammen	0267		0508	0627			
Grafik fehlendes Feedback bei Iteminteraktion	orientierungslos und kein Anhaltspunkt, was stattdessen tun; Mühe gegeben, Humor und Möglichkeiten, die sich aufeinander beziehen	0135	0248	0512 und 0525			0816	
Grafik Grafikfehler	Balken am unteren Rand	0122						
Grafik Grafikmetapher verwirrend	Knöpfe wechseln ihre Lage, weil das Notizbuch einem echten Buch nachempfunden ist; erster Screen überfordert (01)	0124		0414				
Grafik gute Grafikmetapher	Buchecke angedeutet und funktioniert					0709		
Grafik Hintergrundgrafik	ok; Feuer passt nicht zum Rest, nicht schön animiert			0530	0625			
Grafik Items nicht sichtbar am unteren Bildschirmrand	Testperson sieht nicht, was sie gerade im Inventar aufgenommen hat			0409				

Abb. 3. Tabelle mit thematisch geordneten Indizes

Spielenden auslösen, dass beispielsweise ein Lob nach dem erfolgreichen Abschluss eines Minispiels, das zwei Mal in derselben Art und Weise vom Sprecher vorgetragen wird, von den Spielenden nicht als glaubwürdig angenommen wurde. Sie konnten so „Winterfest“ nicht ernst nehmen und wurden sich der Spielsituation bewusst. Andere zu den Valenzen genannte Spielelemente unterstützten einfach die Spielwelt nicht in dem Maße wie es von den Spielenden erwünscht wurde. Hierzu zählt beispielhaft eine unpassende Schriftart in Sans-serif, die nicht die entsprechende Atmosphäre des Spiels wiedergibt.

Ein kritischer Punkt an dieser Stelle war es, während der Auswertung die Segmente, welche die Spielwelt definieren oder die persönliche Beziehung zu dem Spiel prägen, nach entsprechenden Themen zu

ordnen. Hier halfen Tabellen, anhand derer eine Übersicht über die Indizes und die Themenschwerpunkte geschaffen wurde. Eine beispielhafte Auswahl genannter Merkmale ist in Tabelle 1 zu erkennen.

Insgesamt war festzustellen, dass „Winterfest“ von 8 von 9 Testpersonen nicht als schlecht empfanden. Dennoch äußerten 6 von 9 Testpersonen, sie würden es nach der Besprechung nicht freiwillig zu Ende spielen. Die in der vorhergehenden Umfrage herausgestellten, wichtigen Merkmale für Adventure-Spiele (starke Geschichte, viel Witz und Rätsel), konnten von den Testpersonen überwiegend in dem Serious Game wiedergefunden werden. Allerdings identifizierten sie einige Spielsituationen so, dass die Situationen als für Kinder passender wahrgenommen wurden. Weitere gesetzte Valenzen deuteten darauf hin, dass die Geschichte nicht

immersiv genug war, als dass die ungewünschte Assoziation zu Kinderspielen ausgeglichen werden konnte. Die Motivation für weiteres Spielen blieb insgesamt bei fast allen Spielenden aus.

Aus dieser Beobachtung kann geschlossen werden, dass es wichtig ist, die typischen Merkmale des jeweiligen Genres besonders stark in einem Serious Game ausprägen, da dadurch Flow ermöglicht werden kann. Dennoch ist es wichtig, dies für die Zielgruppe adäquaten Rahmen umzusetzen, da ansonsten Unter- oder Überforderung entstehen kann. Hierdurch kann das Spiel von den Spielenden einer unpassenden Zielgruppe zugeordnet werden, wodurch die Lern- und Spielsituation bewusst und Flow unmöglich wird. [Tab. 1]



Kategorie	Name	Beschreibung	Indiz
Design	Schriftart unpassend	Arial passt nicht; wirkt nicht wie von Hand geschrieben	0258, 0502
Geschichte	Unspektakuläre Story	Twist in der Geschichte ist gut, sonst eher mittelspannend; plausibel; weder supraaufregend noch blöd; etwas Abgedrehtes fehlt; wirkt nicht stark genug, nimmt Motivation	0259, 0327, 0442, 0528, 0635, 0923
Minispiel	unglaublicher Lob im Minispiel	Lob kann nicht ernst genommen werden, weil die Aufgabe zu einfach ist; weil 2 Mal hintereinander auf die selbe Art gelobt wird; zu oft gelobt	0158, 0258, 0314, 0329 0630, 0715
Spiel insgesamt	Offensichtlicher Spielverlauf	Keine Items sehr versteckt, nicht viel nachdenken, was als nächstes zu tun ist; im Dialog wird folgender Lernabschnitt zu offensichtlich; untypisch, dass Spieler immer wissen, was sie tun müssen	0259, 0313, 0517, 0620, 0734, 0916
Spiel insgesamt	Spiel für Kinder	Wirkt wie Spiel zum Lernen des Umgangs mit einem PC; einfache Rätsel; andauernde Tutorial-Erklärungen; 100 als Ergebnis simpel und unrealistisch; Lob für Kinder gut, weil die tatsächlich vielleicht noch denken müssen; Kinderrechenaufgabe; „Grundrechnung“; klar was zu tun ist; Sätze zu klar ausformuliert wie eine ‚Heile Welt‘; kein hintergründiger Humor; wie von Eltern für ihre Kinder	0212, 0242, 0245, 0248, 0250, 0254, 0257, 0313, 0402, 0423, 0438, 0503, 0512, 0517, 0532, 0630, 0640, 0715, 0927
Witz	absurder Kommentar	Witzig, abgedreht	0239, 0241, 0245, 0250, 0308, 0311, 0316, 0415, 0420, 0427, 0515, 0616, 0710, 0713, 0817
Witz	zu wenig Humor insgesamt	Humor würde zu mehr Langzeitmotivation führen; zu ernst, zu wenig Ironie; Geschichte kann so bleiben, aber zu wenig Humor; keine unpassenden Elemente wie ‚Schwarze‘ oder so im Spiel; fehlende Item-Kommentare wie bei Monkey Island;	0245, 0262, 0313, 0327, 0644

Tab. 1.
beispielhafte Auswahl von 7 Indizes

Welchen Mehrwert erbrachte der Einsatz von Valenzen in der Studie?

Durch den Einsatz von Valenzen konnten die einzelnen UX-Merkmale, welche für die Spieler in Bezug auf ihre Erwartungen relevant waren, in dem Serious Game markiert werden. Durch diese Dokumentation wurde es möglich, diese UX-Merkmale nach Häufigkeiten auszuwerten, was eine Bewertung vereinfachte. Ebenso konnten die UX-Merkmale sowie ihre übergeordneten Themen qualitativ evaluiert werden. Dies gab Aufschluss darauf, welche UX-Merkmale im Zusammenspiel mit den Erwartungen an das Genre die Entwicklung des Spielflows bestimmten, sowie in welchem Maße gefördert oder gehindert haben.

Literatur

- Burmester, M., Jäger, K., Festl, L. & Mast, M. (2011): Studien zur formativen Evaluation der User Experience mit der Valenzmethode. In S. Schmid et al. (Eds.), Reflexionen und Visionen der Mensch-Maschine-Interaktion – Aus der Vergangenheit lernen, Zukunft gestalten. 9. Berliner Werkstatt Mensch-Maschine-Interaktion, 5. bis 7. Oktober 2011. Fortschritt-Berichte VDI Reihe 22 Nr. 33. (pp. 567–572). Düsseldorf: VDI-Verlag.
- Csikszentmihályi, M. (1992): Flow. Stuttgart: Klett-Cotta.
- de Jong, M.D.T. & van den Haag, M.J. (2002). Exploring Two Methods of Usability Testing: Concurrent versus Retrospective Think-Aloud Protocols. Twente: University of Twente.
- Fritz, J. (1997): Zwischen Transfer und Transformation. Überlegungen zu einem Wirkungsmodell der virtuellen Welt, . In J. Fritz & W. Fehr (Hrsg.), Handbuch Medien. Computerspiele – Theorie, Forschung, Praxis (S. 229–246). Bonn: Bundeszentrale für politische Bildung.
- Hassenzahl, M. (2008): User Experience (UX): Towards an experiential perspective on product quality. In Proc. of IHM'08, 2–5 Sept. 2008, Metz, France, 11–15.

