

Entwicklung und Evaluierung der AR-Applikation „Magneto“

Kurzfassung

Ziel dieses Projektes ist die Entwicklung von Augmented-Reality(AR)-Applikationen. Mit deren Hilfe kann die reale Lernumgebung bzw. das Realexperiment gezielt mit computergenerierten Informationen überblendet werden. Die Sekundarstufe I bietet in der 9. Jahrgangsstufe in Bayern zum Thema der Elektrizitätslehre viele Experimente zur Anwendung einer augmentierten Lernumgebung. Dabei sollen die in diesem Projekt entwickelten Applikationen hauptsächlich die Modelle der magnetischen Felder sichtbar machen. In diesem Beitrag wird der Prozess der Entwicklung und Evaluierung einer AR-Applikation am Beispiel der App „Magneto“ vorgestellt werden. Die Evaluation wird mittels einer Mixed-Methods-Studie durchgeführt. Dabei kommt der quantitative System Usability Score (SUS) nach Brooke und das qualitative Usability Konzept nach Nestler zum Einsatz. Mittels der interaktionsbezogenen Technikaffinität (Affinity for Technology Interaction (ATI)) soll zudem auf einen möglichen Zusammenhang der individuellen Technikaffinität der Probanden und deren Bewertung der Usability bezüglich der App eingegangen werden.

Vorhaben

Die Entwicklung einer AR-Applikation kann als kontinuierlicher Kreisprozess gesehen werden. Zu Beginn steht die Auswahl der passenden Experimente. Dabei ist es von Vorteil, die technische Umsetzung bereits mit einzubeziehen, da nicht jeder Versuch dafür geeignet ist. Ein wichtiger Punkt bei der Auswahl ist z.B. die Geschwindigkeit des Versuchsablaufs, da statische Vorgänge im Allgemeinen besser erkannt werden können. Dann folgen die Programmierung und Erstellung der Applikation. Die Usability spielt hierfür eine große Rolle. Diese Benutzerfreundlichkeit wird in verschiedene Kategorien unterteilt, wie z.B. Nutzen, Intuitivität, Einprägsamkeit, Erlernbarkeit und die persönliche Einstellung. Alle diese Eigenschaften, sollten im Sinne des Nutzers sein, sodass dieser die Applikation akzeptiert und nutzen wird. Um die Usability erheben zu können, gibt es z.B. das Qualitative Usability Konzept. Mit diesem standardisiertem qualitativen Leitfadenterview werden die Probanden befragt. Durch die vielen verschiedenen Kategorien und Antworten der Probanden kann eine anschließende Anpassung der Applikation oder aber auch die Auswahl eines anderen Versuches geschehen. Somit würde man den Kreis erneut durchlaufen.

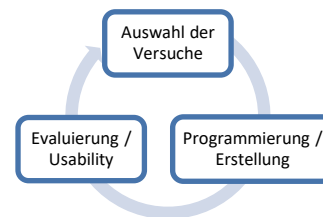


Abb. 1: Durchlaufene Prozesse bei der Entwicklung und Evaluierung der Applikation

Konstrukt der Usability

Bei der Entwicklung der Applikation ist die Usability von entscheidender Bedeutung. Diese wird in der ISO 9241-11 wie folgt definiert: „The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use“ (Bevan et al. 2015, S. 144).

Somit besteht die Usability aus drei Elementen. Das Element der Effektivität bezeichnet die Genauigkeit und Vollständigkeit, mit der ein Benutzer ein bestimmtes Ziel erreichen kann. Die Komponente der Effizienz stellt das Kosten-Nutzen-Verhältnis dar, indem sie den notwendigen Aufwand mit der Genauigkeit und Vollständigkeit der Zielerreichung in Relation setzt. Als dritten Teil wird die Zufriedenheit gesehen, welche eher subjektive Kriterien fokussiert und in der ISO-Norm 9241-11 als die Freiheit von Beeinträchtigungen und als eine allgemein positive Einstellung gegenüber der Systemnutzung definiert ist (Sarodnick und Brau 2016, S. 37).

Für die Bewertung der Effektivität und Effizienz müsste eine Gruppe ohne Hilfsmittel von AR die gestellten Aufgaben bearbeiten und danach eine vergleichende Bewertung der Ergebnisse zu der Treatmentgruppe vollzogen werden. Dieses Vorgehen ist in der jetzigen Phase noch nicht vorgesehen, sodass sich deshalb nur auf die Zufriedenheit der Nutzer konzentriert wird. Diese Zufriedenheit der Nutzer kann sowohl quantitativ mit dem System Usability Scale (kurz: SUS) (Brooke 1996) als auch qualitativ mit dem qualitativen Usability Konzept (Nestler et al. 2011) geschehen. Quantitative Erhebungen bieten durch ihre standardisierte Durchführung einen Vergleich zu den getätigten Aussagen. Qualitative Erhebungsformen der Usability haben den Vorteil, dass diese eine individuelle Sicht der Probanden auf das System zulassen (Nestler et al. 2011). Dieser subjektive Eindruck des Systemnutzers kann jedoch auch einen Einfluss auf die Usability haben (Franke et al. 2019; Wessel 2018). Haben die Probanden ein Problem bei der Verwendung erfahren, generalisieren manche davon dies auf die ganze Applikation. Auf diese Weise werden die Aussagen des qualitativen Usability Konzepts negativ beeinflusst, da durch die Kodierung ein schlechterer Usabilityscore erzielt wird.

Aus diesem Grund wurde die Evaluierung der AR-Applikation in beiden Erhebungsformen durchgeführt und zusätzlich die interaktionsbezogene Technikaffinität erhoben (kurz: ATI – Affinity for Technology Interaction) (Attig et al. 2018). Die Bewertungen der Usability wurde anschließend in Korrelation zur ATI betrachtet. In dieser Phase waren zwei Fragen wesentlich. Zum einen stellte sich die Frage, ob sich das Usability Konzept zur Berechnung eines Usability Scores eignet, welcher mit dem Score aus dem quantitativen SUS übereinstimmt. Zum anderen, ob technisch affine ProbandInnen in beiden Erhebungsformen einen höheren Usability Score für die Applikation vergeben.

Entwicklung der Applikation „Magneto“

Eine genauere Darstellung der Vorgänge wäre zu umfangreich, sodass in diesem Abschnitt nur Eckpunkte angegeben werden. Wie bereits in der Einleitung erwähnt, stellt bei der Auswahl des Versuches die Versuchsablaufgeschwindigkeit das größte Problem dar. Hinzu kommt eine teils mangelhafte Erkennung der Targets, welche die Grundfläche im Raum definiert und die vorliegende Lernumgebung identifiziert. Diese Problematik sollte sich jedoch durch die fortschreitende Entwicklung immer präziserer Kameraobjektive auf der Hardware- und fortlaufenden Updates auf der Softwareseite bald lösen lassen. Für das Entwickeln einer Applikation wurden im Wesentlichen die drei Programme Unity3D, Vuforia und Blender benutzt. Somit ist eine Applikation entstanden, welche dem Nutzenden die Möglichkeit bietet, einen Versuch der sechs verschiedenen Lernstationen zum Thema Magnetismus auszuwählen. Diese Thematik wird in Bayern ab dem Schuljahr 2022/23 in der 10. Klasse behandelt (Staatsinstitut für Schulqualität und Bildungsforschung 2020). Von den Versuchen wurden bis zum jetzigen Zeitpunkt nur die Versuche der Station *Magnetfeldlinien* und der *Versuch von Oersted* evaluiert. Dafür wurden pro Versuch jeweils fünf ProbandInnen akquiriert. Die Evaluierung der anderen Versuchsteile ist in Planung.

Evaluierung der Applikation

Das Zeitfenster bei allen ProbandInnen beträgt 45 Minuten und setzt sich aus vier verschiedenen Abschnitten zusammen.

Einführung (ca. 8 min): Zu Beginn der Datenerhebung erhalten die ProbandInnen einen Fragebogen zu ihrer Person und den Test mit 9 Items für die interaktionsbezogene Technikaffinität. Auf Nachfrage, ob die Thematik bekannt ist, wurde bei Bedarf ein kurzes Video abgespielt. In diesem ist der reale Ablauf des Versuches zu sehen. Neben einer kurzen fachlichen Erklärung sollen die ProbandInnen bei ihrem Wissensstand abgeholt werden und der Fokus auf den bevorstehenden Aufgabenteil gelegt werden.

Arbeitsphase (ca. 15 min): In der darauffolgenden Phase bearbeiten die ProbandInnen eine Station mit den dafür entwickelten Aufgaben. Dabei müssen sie auch die entsprechende Applikation benutzen. In den 15 Minuten dieser Phase bleibt auch genug Zeit, dass die ProbandInnen sich zu Beginn mit der App vertraut machen können.

SUS (ca. 2 min): Mittels des quantitativen System Usability Score wird mit 10 Items ein Usabilitywert erzeugt. Dieser wird mit dem Ergebnis der qualitativen Umfrage ins Verhältnis gesetzt. Aufgrund der kurzen Ausfülldauer wird mit dem Ergebnis eine Tendenz der Benutzerfreundlichkeit erzeugt.

Qualitative Usabilityumfrage (ca. 20 min): Abschließend wird das Interview durchgeführt. mit den 19 Fragen (vgl. Schwanke und Trefzger 2021) sollen Aspekte aufgegriffen werden, welche durch den Kurztest nicht erfasst werden können. Durch die persönliche Befragung kann individuell auf die Probleme der Tester eingegangen werden. Nach Abschluss des Interviews ist die Datenerhebung abgeschlossen.

Vorläufige Ergebnisse

Neben einer Vielzahl von kleinen Anmerkungen und Verbesserungsvorschlägen ist durch das Usabilitykonzept ein Wert erzeugt worden, welcher mit dem Wert aus dem SUS verglichen wird. Dabei konnte festgestellt werden, dass beide Werte miteinander korrelierten (vgl. Abb.2). Zusätzlich erhält man bei der Auswertung eine genaue Aufspaltung der einzelnen Kategorien. Diese Aufspaltung bietet einen Wert pro Haupt- und Unterkategorie. Trägt man diese grafisch auf, so konnten deutliche Unterschiede zwischen den beiden getesteten Applikationen festgestellt werden. So konnte für den Versuch von Oersted durch die Antworten der ProbandInnen der Wert für die Zufriedenheit durch eine mangelnde Interaktion mit dem Versuch erklärt werden. Für die Applikation Magnetfeldlinien hingegen ist klar geworden, dass ein massiver Kritikpunkt die teilweise mangelhafte Erkennung des Targets war. Dies spiegelte sich im Wert für die Zuverlässigkeit im Usabilitykonzept wider. Zwischen ATI und dem erzeugtem Usabilitywert konnte ein mittelstarker linearer Zusammenhang festgestellt werden. Im Zuge dessen reihen sich die vorläufigen Ergebnisse in die Prognosen ein, welche bereits Attig et al. damit erklärten, dass „Personen mit hohem ATI-Score mit mehr Neugier und Exploration versuchen sollten, die Fehlfunktion zu beheben“ (Attig et al. 2018, S. 7).

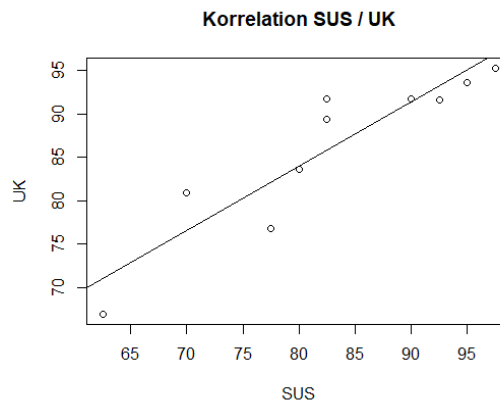


Abb. 2: Aufgetragene Daten des SUS und des Usabilitykonzepts

Literaturverzeichnis

Attig, Christiane; Mach, Sebastian; Wessel, Daniel; Franke, Thomas; Schmalfuß, Franziska; Krems, Josef (2018): Technikaffinität als Ressource für die Arbeit in Industrie 4.0. In: Professur Arbeitswissenschaft und Innovationsmanagement (Hg.): Interact Conference. Innovation der Innovation - Neu Gedacht, Neu Gemacht. Chemnitz, 28.06.2018-29.06.2018. Technische Universität Chemnitz. Band 3.

Bevan, Nigel; Carter, James; Harker, Susan (2015): ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In: Masaaki Kurosu (Hg.): Human-computer interaction. Design and evaluation ; 17th international conference, HCI International 2015, Los Angeles, CA, USA, August 2 - 7, 2015 ; proceedings, part I, Bd. 9169. Cham: Springer (Lecture notes in computer science Information systems and applications, incl. Internet/web, and HCI, 9169). Online verfügbar unter https://link.springer.com/chapter/10.1007%2F978-3-319-20901-2_13.

Brooke, John (1996): SUS - A quick and dirty usability scale. In: Patrick W. Jordan (Hg.): Usability evaluation in industry. Based on the International Seminar Usability Evaluation in Industry that was held at Eindhoven, The Netherlands, on 14 and 15 September 1994]. London: Taylor & Francis, S. 189–194.

Franke, Thomas; Attig, Christiane; Wessel, Daniel (2019): A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. In: *International Journal of Human-Computer Interaction* 35 (6), S. 456–467. DOI: 10.1080/10447318.2018.1456150.

Nestler, Simon; Artinger, Eva; Coskun, Tayfur; Yildirim-Krannig, Yeliz; Schumann, Sandy; Maehler, Mareike et al. (2011): Assessing qualitative usability in life-threatening, time-critical and unstable situations. *GMS Medizinische Informatik, Biometrie und Epidemiologie*; 7(1):Doc01; ISSN 1860-9171 / *GMS Medizinische Informatik, Biometrie und Epidemiologie*; 7(1):Doc01; ISSN 1860-9171. DOI: 10.3205/mibe000115.

Sarodnick, Florian; Brau, Henning (2016): Methoden der Usability Evaluation. Wissenschaftliche Grundlagen und praktische Anwendung. 3., unveränderte Auflage. Bern: Hogrefe. Online verfügbar unter <http://elibrary.hogrefe.de/9783456955971>.

Schwanke, Hagen; Trefzger, Thomas (2021): Entwicklung von AR-Applikationen für die Elektrizitätslehre der Sekundarstufe I. In: H. Grötzebauch und J. Grebe-Ellis (Hg.): *PhyDid B - Didaktik der Physik - Beiträge zur virtuellen DPG-Frühjahrstagung 2021*, S. 421–426. Online verfügbar unter <http://www.phydid.de/index.php/phydid-b/article/view/1166>.

Staatsinstitut für Schulqualität und Bildungsforschung (2020): *LehrplanPLUS. Fachlehrplan - Gymnasium Physik*. Hg. v. Staatsinstitut für Schulqualität und Bildungsforschung. Online verfügbar unter <https://www.lehrplanplus.bayern.de/schulart/gymnasium/fach/physik/inhalt/fachlehrplaene>, zuletzt geprüft am 05.11.2020.

Wessel, Daniel (2018): Warum Technik-Enthusiasten zu positives Feedback geben - und wie man das aufdeckt. Unter Mitarbeit von Rainer Gibbert. Hg. v. www.produktbezogen.de. Online verfügbar unter <https://www.produktbezogen.de/wie-man-zu-positives-feedback-aufdeckt/>.