


Minimal-invasive Messung lernrelevanter Parameter für den Einsatz im Game-based Learning

Dietmar Zoerner, Paul Beschorner, Lars Michel, Ulrike Lucke  ¹

Abstract: Für die zielgerichtete Entwicklung und Nutzung von Lernanwendungen ist es nötig nachzuvollziehen, welche Ursache-Wirkungs-Beziehungen zum Entstehen von Lernergebnissen beitragen. Bei Lernspielen ist dies besonders schwierig, weil klassische Erhebungsinstrumente (wie z.B. Fragebögen) das Spielerleben und damit den Lerneffekt beeinträchtigen können. Der Beitrag stellt anhand von Fallstudien zwei Ansätze zur Messung von Aufmerksamkeit und Flow als lerneffekt-vermittelnde Parameter direkt in der Lernanwendung vor, analysiert die damit erfassbaren Daten und diskutiert die Möglichkeiten und Grenzen derartiger Erhebungsmethoden.

Keywords: Game-based Learning, Evaluation, Flow, Aufmerksamkeit.

1 Evaluationsdesigns im Game-based Learning

Game-based Learning [Pr01] nutzt spielerisch entwickelte Kompetenzen zur Erreichung formeller Bildungsziele. Die Analyse von Ursache-Wirkungs-Beziehungen und erreichbaren Effekten ist in den oft komplexen Arrangements (mit z.B. hohen Freiheitsgraden oder adaptiven Settings) schwierig, da eine Isolation einzelner Parameter nur unter Laborbedingungen möglich ist, was aber die Validität bzw. Übertragbarkeit der Erkenntnisse unter Realbedingungen offen lässt [Pe19]. Dieser Beitrag untersucht, wie lernrelevante Parameter während eines Lernspiels erhoben werden können, ohne das Spielerlebnis und damit den Lerneffekt zu beeinträchtigen. Dabei wird auf Aufmerksamkeit [HL79] und Flow [Cs14] als Vermittler von Lernerfolg fokussiert.

Zur Erfassung von Lernerfolgen und deren Zustandekommen werden i.d.R. Fragebögen, Beobachtungen oder Messungen eingesetzt. Es existieren z.B. Fragebögen zu User Experience und Usability [THS18] sowie für spielspezifische Aspekte [JGP18]. Fragebögen stoßen jedoch an Grenzen, da sie nur subjektive Momentaufnahmen aus Sicht der Befragten liefern. Zudem ermöglichen Fragebögen nach der Nutzung der Anwendung nur einen Überblick über die komplette Spieldauer. Kurze Umfragen im Verlauf des Spiels können dem entgegenwirken und ermöglichen eine Auswertung der Parameter über den Verlauf des Spiels. Sie unterbrechen jedoch den Spielfluss, können die Konzentration und den Flow der Probanden stören und die Messung verfälschen. Auch Beobachtungen zielen nicht auf die intern wirkenden Mechanismen ab.

¹ Universität Potsdam, Institut für Informatik und Computational Science, An der Bahn 2, 14476 Potsdam, firstname.lastname@uni-potsdam.de, <https://orcid.org/0000-0003-4049-8088>

Es bleiben Messungen [ME18] als Ausgleich zwischen ökologischer Validität und experimenteller Kontrolle. Eine Möglichkeit sind physiologische Parameter, wie z.B. Puls oder Hautleitfähigkeit [NDG10]. Diese Messmethoden unterbrechen den Spielfluss nicht und liefern kontinuierlich Daten. Allerdings können die Sensoren bzw. Labor-Settings die Probanden stören bzw. ablenken. Alternativ können die Daten direkt in der Anwendungs-Software z.B. über Tastatureingaben, Mausbewegungen, Reaktionszeiten [Os16] oder Leistungen in der Anwendung erfasst werden. In diesem Beitrag wird die indirekte Messung von Aufmerksamkeit und Flow verfolgt.

2 Messung von Aufmerksamkeitsparametern

Das hier vorgestellte Mess-Szenario ist auf die Untersuchung der Aufmerksamkeit von Lernenden [HL79] während des Spielens, insbesondere während des Wechsels zwischen verschiedenen Endgeräten gerichtet, am Beispiel des Trainings sozioemotionaler Kompetenzen insbesondere durch autistische Nutzende [ZML17]. Für das Lernsystem Lodur wurde das verbreitete Spiel Minecraft um Narrationen und das synchronisierte Bespielen einer speziellen mobilen App erweitert [St20], siehe Abb. 1.



Abb.

1: Lodur verbindet eine Minecraft-Anwendung mit einer mobilen App (links).
Ein Geschicklichkeitsspiel ist in Minecraft als Element der Messumgebung eingebettet (rechts).

Der Wechsel zwischen Desktop-Anwendung und mobiler App könnte das immersive Spielerlebnis beeinträchtigen und zu negativen Auswirkungen für die Aufmerksamkeit und den Trainingseffekt führen, was zu untersuchen ist. Die Messung von Indikatoren des Aufmerksamkeitsniveaus ist sensibel für Einflüsse durch das Messverfahren selbst. Daher wurde ein spielbasiertes Messverfahren entwickelt, das diese Indikatoren erhebt, ohne die Probanden in ihrer Aufmerksamkeit zu stören. Das Mess-Szenario basiert auf einem Geschicklichkeitsspiel, bei dem mehrere Parameter durch das Aufmerksamkeitsniveau der Spielenden beeinflusst werden. In mehreren Durchläufen sollen Tiere von einem Laufband geschossen werden; siehe Abb. 1 (rechts). Zielgeschwindigkeit, Zielgenauigkeit und Fehlerhäufigkeit dienen als Indikatoren für das Aufmerksamkeitsniveau. Zwischen einigen Übungen werden kurze Aufgaben mit der mobilen App absolviert. Pandemiebedingt nahmen nur elf Probanden teil. Die Messwerte eines Spieldurchlaufes sind aufgrund der hohen individuellen Varianz nicht sinnvoll mit einem absoluten Wert vergleichbar und werden daher pro Proband verglichen. Die Leistung der Probanden mit und ohne Nutzung der App zeigt Abb. 2 links; dabei wird der Leistungswert errechnet als

$(10000 \cdot \text{bonus})/t$ mit $\text{bonus}=1$ bei Fehlschüssen und $\text{bonus}=1.25$ sonst sowie t = Spielzeit des Levels in ms. Die Messwerte zeigen eine Streuung zwischen den Individuen, aber keine wesentliche Abweichung in Abhängigkeit von der vorherigen App-Nutzung.

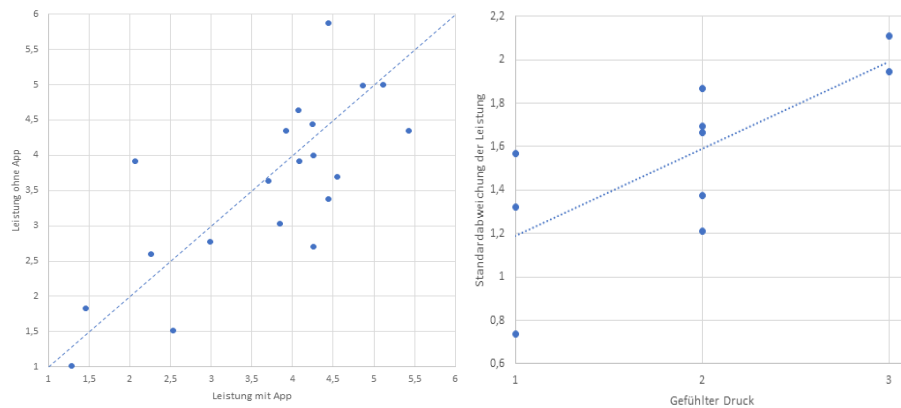


Abb. 2: Auswertung der Störung durch die Nutzung der mobilen App auf die Leistung

Für eine aussagekräftige Messung ist die Schwierigkeit der Aufgaben relevant. Die Leistungsfähigkeit der einzelnen Probanden hängt von individuellen Fähigkeiten und Trainingsständen ab. Jedoch ist die Stärke der Leistungsschwankungen der Spielenden ein Indikator für deren Aufmerksamkeit. Um die Qualität des Verfahrens einordnen zu können, wurde zusätzlich nach dem Ende der Messung durch Fragebögen das subjektive Erleben der Anwendung erhoben. Es zeigt sich eine Korrelation zwischen Streuung der Leistungswerte und subjektivem Leistungsdruck, wie in Abb. 2 rechts gezeigt. Dies könnte damit zu erklären sein, dass ein hoher Leistungsdruck das Flow-Erleben stört [Cs14] und die Leistungsfähigkeit beeinträchtigt. Der Pearson-Korrelationskoeffizient 0,74 ist für die geringe Anzahl der Proband*innen gut und belegt die grundlegende Funktionsfähigkeit des minimal- invasiven Messverfahrens.

3 Messung von Flow-Parametern

Das hier vorgestellte Evaluationsdesign ist auf die Untersuchung von Flow [Cs14] gerichtet. GraffitiVR (siehe Abb. 3) ist eine Laborstudie zu Virtual Reality für das Lernen in Anlehnung an die VR-Lackierwerkstatt [Ze20]. Auf einer industriellen Map wurden Plakatwände und Farbeimer (links) sowie eine Sprühpistole (rechts) platziert. Die Sprühfarbe kann über die Eimer ausgewählt werden, um die Plakatwand zu besprühen. Der Kern der Anwendung bestand in einer unmittelbaren Messung des Flow-Erlebens [Cs14]. Für VR-Lernanwendungen werden zwei Aspekte des Flows gemessen: tiefe Konzentration (Aufmerksamkeit auf die Aufgabe, nicht andere Informationen) und verändertes Zeitgefühl (Fokus auf die Gegenwart).



Abb. 3: In GraffitiVR wird ein virtueller Hinterhof zum Besprühen einer Wand genutzt (links). Dort wird mit einer Sprühpistole in verschiedenen Farben eine Plakatwand besprüht (rechts).

Die Messung der Konzentration erfolgt durch Blickrichtung über die normale Durchführung der Anwendung, also dem Fokus auf das Sprayen der Plakatwand. Währenddessen sollen visuell-akustische Ablenkungen die Aufmerksamkeit stören. Für Flow spricht, wenn Ablenkungen nicht wahrgenommen werden. Abb. 4 zeigt eine beispielhafte Messreihe der Konzentration; Ablenkungen sind als Farbbalken hinterlegt.

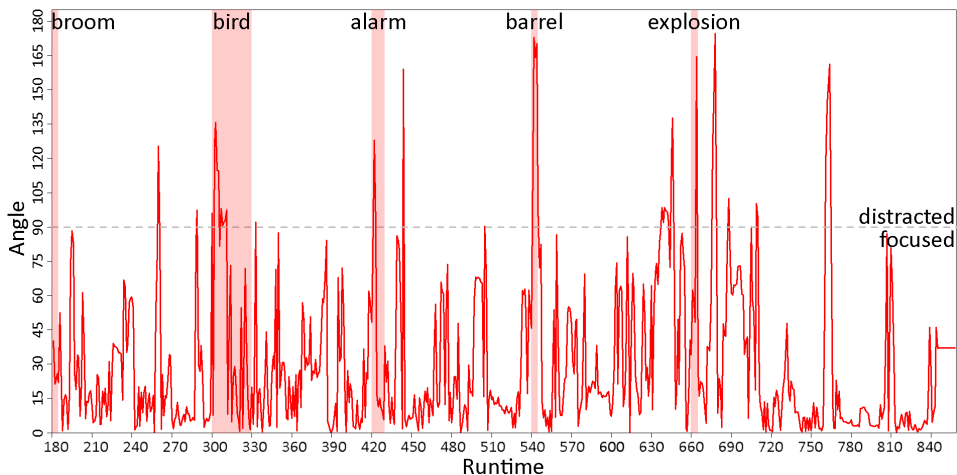


Abb. 4: Veränderung der Blickrichtung von Teilnehmer C (Winkel) über die Zeit. Die Plakatwand befindet sich vor dem Teilnehmer. Der Blick nach hinten ($>90^\circ$) wird als Ablenkung gewertet.

Das Zeitgefühl wurde durch die Aufgabe gemessen, einen Farbeimer regelmäßig zu schütteln. Weitere Datenerhebungen erfolgten durch zwei Fragebögen zum Vergleich mit den Messungen in VR: dem Swedish Flow Proneness Questionnaire [UMA12] zur Flow-Neigung im Alltag sowie der Flow-Kurzskala [RVE03] zum Flow-Erleben. Zwischen diesen Fragebögen erfolgte die Durchführung der Anwendung für mindestens zwölf Minuten (damit alle fünf Ablenkungen auftreten können). Pandemie-bedingt nahmen nur drei Forschende an der Laborstudie teil. Es sind keine Muster in den Daten erkennbar, wohl aber individuelle Unterschiede. So hat etwa Proband A nur eine Ablenkung wahrgenommen, aber trotzdem den Farbeimer regelmäßig geschüttelt. Proband C hat alle Ablenkungen bemerkt, jedoch das Schütteln des Eimers (nach dem ersten Mal) vollkommen vergessen. Dennoch zeigten beide Probanden relativ hohe Werte bei der Auswertung ihres Flow-Empfindens. Proband B zeigte mit wenigen wahrgenommenen

Ablenkungen und einem deutlichen Abstand zwischen den Eimerschütteln ein starkes Flow-Erleben. Deutliche Widersprüche zwischen Fragebögen und Messwerten konnten also nicht identifiziert werden, aber auch keine Korrelationen.

4 Zusammenfassung und Ausblick

Beide Studien zeigen die Machbarkeit minimal-invasiver Messungen von Aufmerksamkeit und Flow. Für ausgewählte Aspekte in etablierten Fragebögen konnten Entsprechungen in der Software bzw. im Nutzerverhalten gefunden werden, und die erhobenen Daten zeigen keine Widersprüche zwischen den objektiv gemessenen und den subjektiv erfragten Aspekten. Dennoch sind bislang keine absoluten Flow- oder Aufmerksamkeitswerte ableitbar, sondern nur relative intrasubjektive Aussagen. Es bleibt offen, ob sich hinreichend viele und aussagekräftige Indikatoren finden lassen, die eine intersubjektiv vergleichbare Messung von Aufmerksamkeit bzw. Flow erlauben. Für Flow erscheint dies angesichts des präzisen Modells und der konsistenten Fragebögen derzeit leichter möglich als für Aufmerksamkeit, die als komplexes Konstrukt mit verschiedenen Dimensionen und Erhebungsverfahren nicht einheitlich beschrieben ist.

Interessant ist die Verschiedenartigkeit der Messwerte. Es wurden kontinuierliche Skalen (z.B. Trefferquote beim Schießen) und binäre Aussagen (z.B. Störung wahrgenommen) verwendet. Der Abgleich mit Likert-Skalen stellt gleichermaßen eine Herausforderung und eine Chance für differenzierte Erhebungsmethoden dar. In beiden Studien war die Schwierigkeit u.U. zu niedrig, was durch Ansätze von Adaptivität vermeidbar wäre. Die entworfenen Messelemente sind kein originärer Teil einer Lernanwendung, sondern wurden künstlich hinzugefügt. Für die Verifikation der Verfahren ist das sinnvoll. Die Experimentierumgebung erlaubt eine Kontrolle der untersuchten Anwendungen trotz variierender Kontexte. Für den Einsatz in Lernanwendungen sind jedoch generische Konzepte erforderlich, die sensibel für Aufmerksamkeit bzw. Flow und zugleich thematisch flexibel einpassbar sind.

Konzeption und Validierung derartiger Messverfahren werden in Informatik und angrenzenden Disziplinen weiter verfolgt. Die Weiterentwicklung von Modellen und Fragebögen in der Kognitionspsychologie wird die Entwicklung minimal-invasiver Messverfahren in Lernanwendungen befördern. Insbesondere die systematische Ableitung messbarer Parameter aus theoretischen Konstrukten erfordert interdisziplinäre Anstrengungen. Dabei müssen die hier nur ausschnittsweise berücksichtigten Parameter von Flow und Aufmerksamkeit um weitere Aspekte ergänzt werden. Zudem sollten auch andere lernrelevante Parameter und deren Zusammenhänge beforscht werden. Besonders herausfordernd und relevant sind kollaborative Szenarien, die bislang in den lernpsychologischen Grundlagen noch kaum modelliert sind.

Danksagung

Teile der hier beschriebenen Arbeiten wurden vom BMBF in den Verbundprojekten Emotisk (2015-2018, 16SV7241) und HandLeVR (2019-2021, 01PV18002A) sowie von der Mathematisch-Naturwissenschaftlichen Fakultät der Uni Potsdam im Programm zur Chancengleichheit von Nachwuchswissenschaftler:innen (2021) gefördert.

Referenzen

- [Cs14] Csikszentmihalyi, M.: Applications of Flow in Human Development and Education, Springer, 2014.
- [HL79] Hale, G.A.; Lewis, M.: Attention and Cognitive Development. Springer, 1979.
- [JGP18] Johnson, D.; Gardner, M.J.; Perry, R.: Validation of two game experience scales. *International Journal of Human-Computer Studies*, 118, 2018, 38-46.
- [ME18] Molnar, A.; Estrada, J.G.: A Comparative Study of In-Game and Out-Game Assessment for Storyline-Based Games. In: Proc. ICALT 2018.
- [NDG10] Nacke, L.E.; Drachen, A.; Göbel, S.: Methods for Evaluating Gameplay Experience in a Serious Gaming Context, *Int. J. Comput. Sci. Sport*, 9/2, 2010.
- [Os16] Osman, Z. M.; Dupire, J.; Mader, S.; Cubaud, P.; Natkin, S.: Monitoring player attention: A non-invasive measurement method applied to serious games. *Entertainment Computing*, Volume 14, S. 33-43, 2016.
- [Pe19] Peeters, D.: Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review* 26, 2019, S. 894-900.
- [Pr01] Prensky, M.: Digital Game-Based Learning. McGraw-Hill, 2001.
- [RVE03] Rheinberg, F., Vollmeyer, R., Engeser, S.: Die Erfassung des Flow-Erlebens. In: Diagnostik von Motivation und Selbstkonzept, Hogrefe, 2003, S. 261-279.
- [St20] Strickroth, S.; Zoerner, D.; Moebert, T.; Morgiel, A.; Lucke, U.: Game-based promotion of motivation and attention for socio-emotional training in autism. *i-com*, 19/1, 2020, S. 17-30.
- [THS18] Thomaschewski, J., Hinderks, A. & Schrepp, M.: Welcher UX-Fragebogen passt zu meinem Produkt?. In: Proc. MuC 2018 - Usability Professionals, S.437-446.
- [UMA12] Ullén, F., de Manzano, Ö., Almeida, R. et al.: Proneness for Psychological Flow in Everyday Life: Associations with Personality and Intelligence, *Personality and Individual Differences*, 52(2): 167–72, 2012.
- [Ze20] Zender, R.; Sander, P., Weise, M.; Mulders, M., Lucke, U.; Kerres, M.: Action-oriented Learning in a VR Painting Simulator. Proc. 4th Int. Symposium on Emerging Technologies for Education, LNCS 11984, S. 46-51, Springer, 2020.
- [ZML17] Zoerner, D., Moebert T., Lucke U.: IT-gestütztes Training sozio-emotionaler Kognition für Menschen mit Autismus, *Informatik Spektrum* 40/02, S. 546-555, Springer, 2017.