

Wirkungen des Flow-Erlebens in einem Serious Game auf die Veränderung von Fachwissen und Selbstwirksamkeitserwartung

Dissertation zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

Von Phoebe Perlwitz

Aus Heilbronn



Wirkungen des Flow-Erlebens in einem Serious Game auf die Veränderung von Fachwissen und Selbstwirksamkeitserwartung © 2023 by Phoebe Perlwitz is licensed under [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

<https://orcid.org/0009-0006-4519-5501>

<https://doi.org/10.60530/opus-3265>

Promotionsfach: Technik

Erstgutachterin: Prof.in Dr. Jennifer Stemmann

Zweitgutachterin: Prof.in Dr. Katja Scharenberg

Tag der mündlichen Prüfung: 26.01.2024

Zusammenfassung

Die Förderung von Mädchen in Mathematik, Informatik, Naturwissenschaften und Technik, den MINT-Fächern, ist von großer Bedeutung, um ihnen eine berufliche Perspektive in diesem Bereich zu ermöglichen und Geschlechterungleichheiten abzubauen. Trotz erheblicher Bemühungen besteht eine anhaltende Geschlechterkluft in der MINT-Bildung und -Arbeitswelt. Ein Grund hierfür wird in Unterschieden der Selbstwahrnehmung gesehen. Bei objektiv gleichen Leistungen schätzen Schülerinnen ihre Fähigkeiten im MINT-Bereich in Bezug auf zukünftige Erfolge, die Selbstwirksamkeitserwartung, geringer ein als Schüler. Daraus resultiert eine Nicht-Wahl von Fächern und Berufen in diesen Bereichen.

Serious Games – digitale Spiele, die speziell für Bildungs- und Lernzwecke entwickelt werden – könnten das Potenzial haben, die Selbstwirksamkeitserwartung der Schülerinnen zu erhöhen und ermöglichen eine individuelle Unterstützung, was auch für das Lernen förderlich ist. Durch die interaktive und spielerische Natur von Serious Games kann das sogenannte Flow-Erleben ermöglicht werden, einen Zustand völliger Vertiefung und konzentrierter Hingabe an eine Aktivität. Dieses Flow-Erleben kann dazu beitragen, die Selbstwirksamkeitserwartung der Schülerinnen zu stärken, indem es ihnen ermöglicht, ihre Fähigkeiten in einem unterstützenden und engagierten Umfeld zu erkunden und zu entwickeln.

Diese Annahme bildet den Kern dieser Arbeit, welche die Auswirkungen eines Serious Games im Bereich der Elektrizitätslehre auf das Flow-Erleben, das Lernen und die Selbstwirksamkeitserwartung von Schüler:innen untersucht. Insgesamt nahmen 363 Schüler:innen (310 vollständige Fragebögen) an einem Pre-Post-Design mit Kontrollgruppe teil. Die Ergebnisse zeigen, dass das Serious Game in geringer Effektstärke mehr Flow-Erleben ermöglicht als die konventionellen Unterrichtsmaterialien der Kontrollgruppe. Innerhalb der Experimental- und der Kontrollgruppe zeigten sich geschlechtsspezifische Unterschiede: In beiden Gruppen erlebten die Schüler mehr Flow als die Schülerinnen. Das Flow-Erleben war in beiden Gruppen ein Prädiktor für die Selbstwirksamkeitserwartung nach der Intervention, nur in der Kontrollgruppe konnte das Flow-Erleben auch das Fachwissen vorhersagen. Die Studie zeigt auch, dass das Fachwissen vor der Intervention für die Lernwirksamkeit des Spiels nicht von Bedeutung ist.

Die Ergebnisse zeigen Hinweise darauf, dass hohes Flow-Erleben und hohe Selbstwirksamkeitserwartung gemeinsam auftreten, da sie durch ähnliche Komponenten der Settings hervorgerufen werden. Damit könnte die Selbstwirksamkeitserwartung durch geeignete Unterrichtsszenarien verbessert werden, die sowohl die Theorie der Selbstwirksamkeitserwartung als auch Elemente des Flow-Erlebens berücksichtigen. Eine positive Lernumgebung, die die Interessen und Bedürfnisse der Mädchen berücksichtigt, könnte einen wesentlichen Beitrag zur Förderung von Mädchen im MINT-Bereich leisten, um die Geschlechterkluft in der MINT-Bildung und -Arbeitswelt zu überbrücken.

Abstract

Promoting girls in mathematics, computer science, natural sciences and technology, the STEM subjects, is of great importance in order to give them career prospects in this field and reduce gender inequality. Despite considerable efforts, there is a persistent gender gap in STEM education and the world of work. One reason for this is seen in differences in self-perception. Despite objectively equal performance, female students rate their abilities in STEM fields lower than male students in terms of future success, i.e. self-efficacy expectations. This results in them not choosing subjects and professions in these areas.

Serious games - digital games that are specially developed for educational and learning purposes - could have the potential to increase students' self-efficacy expectations and enable individual support, which is also beneficial for learning. The interactive and playful nature of serious games can enable the so-called flow experience, a state of complete immersion and concentrated dedication to an activity. This flow experience can help to strengthen students' sense of self-efficacy by allowing them to explore and develop their skills in a supportive and engaging environment.

This assumption forms the core of this study, which examines the effects of a serious game in the field of electricity on the flow experience, learning and self-efficacy expectations of students. A total of 363 students (310 complete questionnaires) took part in a pre-post design with a control group. The results show that the serious game enables more flow experience than the conventional teaching materials of the control group with a low effect size. There were gender-specific differences within the experimental and control groups: In both groups, male students experienced more flow than female students. In both groups, the experience of flow was a predictor of the expectation of self-efficacy after the intervention; only in the control group was the experience of flow also able to predict subject knowledge. The study also shows that expertise before the intervention is not important for the learning effectiveness of the game.

The results show evidence that high flow experience and high self-efficacy expectancy occur together, as they are caused by similar components of the settings. Thus, self-efficacy expectations could be improved by suitable teaching scenarios that take into account both the theory of self-efficacy expectations and elements of the flow experience. A positive learning environment that takes girls' interests and needs into account could make a significant contribution to promoting girls in STEM fields in order to bridge the gender gap in STEM education and the world of work.

Inhalt

1	Einleitung.....	1
2	Serious Games in der (MINT)-Bildung.....	4
2.1	Serious Games – Spiele mit einem Bildungszweck.....	4
2.1.1	Lernwirksamkeit von Serious Games.....	5
2.1.2	Empirische Untersuchung der Lernwirksamkeit in Serious Games.....	6
2.1.3	Lernförderliche Elemente in Serious Games.....	8
2.1.4	Serious Games im Unterricht und ihre Auswirkungen auf die Lernmotivation.....	9
2.2	Die Rolle des Flow-Erlebens in Serious Games.....	10
2.2.1	Flow-Erleben und Lernen.....	12
2.2.2	Merkmale in Serious Games für ein Flow-Erleben.....	14
2.3	Wirkungen der Serious Games auf die Selbstwirksamkeitserwartung.....	15
2.3.1	Die vier Quellen der Selbstwirksamkeitserwartung in Serious Games als Forschungsgegenstand.....	16
2.3.2	Empirische Befunde für die Veränderung der Selbstwirksamkeitserwartung durch Serious Games	17
2.3.3	Selbstwirksamkeitserwartung und Flow-Erleben.....	19
2.4	Die Rolle des Selbstkonzepts in Zusammenhang mit Serious Games.....	21
2.5	Einordnung der Konstrukte in ein nomologisches Netzwerk.....	23
3	Fragestellungen und Hypothesen.....	26
4	Methodenteil.....	31
4.1	Instrumente.....	31
4.1.1	Fachwissenschaftliche Inhalte der Intervention.....	31
4.1.2	Das Serious Game Lights Out.....	34
4.1.3	Materialien der Kontrollgruppe.....	41
4.1.4	Fachwissenstest.....	41
4.1.5	Erfassung des Flow-Erlebens.....	44
4.1.6	Selbstwirksamkeitserwartung.....	46
4.1.7	Physikbezogenes und computerbezogenes Selbstkonzept.....	47
4.1.8	Weitere Variablen.....	48
4.1.9	Zusammenstellung der Instrumente.....	49
4.2	Statistische Methoden.....	49
4.2.1	Ermittlung von Unterschieden zwischen zwei Gruppen oder Zeitpunkten.....	49
4.2.2	Vorhersage durch Regressionen.....	50
4.2.3	Untersuchungen mit Strukturgleichungsmodellierung.....	53

4.2.4	Verwendete Software und Bibliotheken	55
4.2.5	Statistische Power und Stichprobengröße	55
4.3	Studiendesign und Durchführung	56
5	Ergebnisse.....	57
5.1	Ergebnisse der Pilotierung	57
5.1.1	Pilotierung des Fachwissenstests	58
5.1.2	Pilotierung der Selbstwirksamkeitserwartung	59
5.1.3	Pilotierung des Flow-Erlebens	59
5.1.4	Pilotierung des computerbezogenen und physikbezogenen Selbstkonzepts	59
5.2	Ergebnisse der Hauptstudie	60
5.2.1	Fachwissen.....	61
5.2.2	Selbstwirksamkeitserwartung	62
5.2.3	Flow-Erleben.....	63
5.2.4	Ergebnisse des computerbezogenen und dem physikbezogenen Selbstkonzept.....	63
5.2.5	Unterschiede nach Gruppen und Geschlechtern	64
5.2.6	Ermittlung der Einflussgrößen auf das Erleben von Flow.....	75
5.2.7	Wirkungen des Flow-Erlebens	82
6	Diskussion der Ergebnisse	93
6.1	Pretest.....	93
6.2	Intervention	94
6.3	Posttest	96
6.3.1	Zusammenhänge im Strukturgleichungsmodell der Experimentalgruppe.....	97
6.3.2	Zusammenhänge im Strukturgleichungsmodell der Kontrollgruppe	98
6.4	Limitation und Ausblick.....	100
7	Fazit	103
	Literaturverzeichnis.....	105
8	Anhang.....	1

Abbildungsverzeichnis

Abbildung 1 <i>Marsh-Shavelson-Modell des Selbstkonzepts (übersetzt nach Marsh et al., 1988)</i>	21
Abbildung 2 <i>Nomologisches Netzwerk der theoretischen Zusammenhänge</i>	25
Abbildung 3 <i>Darstellung eines einfachen Schaltkreises</i>	31
Abbildung 4 <i>Kirchhoffsche Knotenregel</i>	32
Abbildung 5 <i>Kirchhoffsche Maschenregel</i>	32
Abbildung 6 <i>links: Messung der Stromstärke, rechts: Messung der Spannung</i>	32
Abbildung 7 <i>Erstellung des Avatars</i>	35
Abbildung 8 <i>Im Haus können unterschiedliche Räume angesteuert werden</i>	36
Abbildung 9 <i>Lösen der Aufgaben im Kinderzimmer</i>	37
Abbildung 10 <i>Buch zur Elektrizitätslehre</i>	37
Abbildung 11 <i>Experimente auf dem Dachboden</i>	38
Abbildung 12 <i>Badges für erfolgreich bearbeitete Inhalte</i>	38
Abbildung 13 <i>Materialien der Kontrollgruppe</i>	41
Abbildung 14 <i>Beispielaufgabe aus dem Fachwissenstest nach Urban-Woldron Hopf (2012)</i>	43
Abbildung 15 <i>Ein Pop-up-Fenster zur Abfrage des Flow-Erlebens im Serious Game</i>	46
Abbildung 16 <i>Pfaddiagramm der Moderation</i>	52
Abbildung 17 <i>Pfaddiagramm der Mediation</i>	53
Abbildung 18 <i>Ergebnisse im Fachwissenstest vor und nach der Pilotierung</i>	58
Abbildung 19 <i>Ergebnisse des Fachwissenstests beider Interventionsgruppen für den A- und B-Test</i> ... 59	59
Abbildung 20 <i>Gegenüberstellung der Pretest und Posttest Ergebnisse</i>	61
Abbildung 21 <i>Gegenüberstellung der Gruppen bei den Fachwissensergebnissen</i>	62
Abbildung 22 <i>Pfaddiagramm der Experimentalgruppe</i>	85
Abbildung 23 <i>Pfaddiagramm der Kontrollgruppe</i>	88
Abbildung 24 <i>Johnson-Neyman-Plot des Moderatoreffekts von Flow auf das Fachwissen</i>	90
Abbildung 25 <i>Nomologisches Netzwerk für die Ergebnisse der Experimentalgruppe</i>	97
Abbildung 26 <i>Nomologisches Netzwerk für die Ergebnisse der Kontrollgruppe</i>	100

Tabellenverzeichnis

Tabelle 1 <i>Die drei Formen der Spielforschung, übersetzt aus R. Mayer (2019).</i>	7
Tabelle 2 <i>Komponenten des Flow-Erlebens ergänzt nach Csikszentmihalyi (2014).</i>	11
Tabelle 3 <i>Die vier Quellen der Selbstwirksamkeitserwartung nach Bandura (1977).</i>	16
Tabelle 4 <i>Hypothesen zur ersten Fragestellung.</i>	28
Tabelle 5 <i>Hypothesen zur zweiten Fragestellung.</i>	29
Tabelle 6 <i>Hypothesen zur dritten Fragestellung.</i>	30
Tabelle 7 <i>Anforderungsbereiche des Fachwissens (KMK, 2004).</i>	42
Tabelle 8 <i>Items der Flow-Kurz-Skala nach Rheinberg et al. (2019).</i>	45
Tabelle 9 <i>Items zur Erfassung der physikbezogenen Selbstwirksamkeitserwartung (verändert nach Jerusalem & Satow, 2009).</i>	47
Tabelle 10 <i>Items zur Erfassung des physikbezogenen Selbstkonzept (verändert nach Weißnigg, 2012).</i>	48
Tabelle 11 <i>Items zur Erfassung des computerbezogenen Selbstkonzepts (Janneck et al., 2012).</i>	48
Tabelle 12 <i>Cutoff-Werte für Strukturgleichungsmodelle.</i>	54
Tabelle 13 <i>Untersuchungsdesign.</i>	56
Tabelle 14 <i>Ergebnisse der Pilotierung.</i>	57
Tabelle 15 <i>Ergebnisse der Erhebung.</i>	61
Tabelle 16 <i>Ergebnisse zur Selbstwirksamkeitserwartung</i>	62
Tabelle 17 <i>Ergebnisse zum Flow-Erleben</i>	63
Tabelle 18 <i>Ergebnisse des physikbezogenen Selbstkonzepts</i>	64
Tabelle 19 <i>Ergebnisse des computerbezogenen Selbstkonzepts</i>	64
Tabelle 20 <i>Prüfung der Ergebnisse auf Normalverteilung</i>	66
Tabelle 21 <i>Vergleich der Schulformen zu zwei Zeitpunkten.</i>	67
Tabelle 22 <i>Vergleich der Ergebnisse der Subgruppen.</i>	68
Tabelle 23 <i>Vergleich der Interventionsgruppen zu zwei Zeitpunkten.</i>	69
Tabelle 24 <i>Vergleich der Interventionsgruppen.</i>	70
Tabelle 25 <i>Geschlechtervergleich in der Experimentalgruppe.</i>	71
Tabelle 26 <i>Geschlechtervergleich innerhalb der Kontrollgruppe.</i>	72
Tabelle 27 <i>Vergleich der Geschlechter in der Experimental- und Kontrollgruppe zu zwei Zeitpunkten.</i>	73
Tabelle 28 <i>Verworfen und verifizierte Hypothesen zur ersten Fragestellung</i>	74
Tabelle 29 <i>Ergebnisse zur Voruntersuchung der Variablen</i>	76
Tabelle 30 <i>Ergebnisse zu den Korrelationen der Variablen.</i>	77
Tabelle 31 <i>Regressionstabelle Flow-Erleben über alle Proban:innen.</i>	78
Tabelle 32 <i>Regressionstabelle Flow-Erleben in der Experimentalgruppe</i>	79
Tabelle 33 <i>Regressionstabelle Flow-Erleben in der Kontrollgruppe.</i>	80
Tabelle 34 <i>Verworfen und verifizierte Hypothesen zur zweiten Fragestellung</i>	81
Tabelle 35. <i>Messmodell des Flow-Erlebens in der Experimentalgruppe</i>	83
Tabelle 36 <i>Regressionstabelle des Strukturgleichungsmodells für die Experimentalgruppe.</i>	84
Tabelle 37. <i>Messmodell des Flow-Erlebens in der Kontrollgruppe</i>	86
Tabelle 38 <i>Regressionstabelle des Strukturgleichungsmodells für die Kontrollgruppe</i>	87
Tabelle 39 <i>Moderatoranalyse von Flow und Fachwissen.</i>	90
Tabelle 40 <i>Verworfen und verifizierte Hypothesen zur dritten Fragestellung</i>	92

1 Einleitung

In einer Welt, die zunehmend von Technologie und Innovation angetrieben wird, ist die Förderung von Schülerinnen in den MINT-Fächern von großer Bedeutung, um ihnen eine Perspektive in diesem Bereich zu ermöglichen und Geschlechterungleichheiten abzubauen. Trotz erheblicher Anstrengungen in den letzten Jahrzehnten, Frauen in MINT-Fächern zu unterstützen und ihre Präsenz in diesen Bereichen zu erhöhen, bleibt die Geschlechterkluft in der MINT-Bildung und -Arbeitswelt bestehen (z. B. Jann & Hupka-Brunner, 2020). Forschungen zeigen, dass Schülerinnen in Bezug auf das Fachwissen den Schülern oft auf Augenhöhe sind, so haben die Programme für internationale Schülerbewertung (PISA) gezeigt, dass es keine signifikanten geschlechtsspezifischen Unterschiede in den MINT-Leistungen gibt (OECD, 2009, 2016). Dennoch ist bei Schülerinnen im Vergleich zu Schülern häufig eine geringere Selbstwirksamkeitserwartung die Überzeugung einer Person, bestimmte Handlungen auszuführen und gewünschte Ergebnisse erzielen zu können (Bandura, 1977) und Motivation in den MINT-Fächern festzustellen (Solga & Pfahl, 2009). In der Folge haben Frauen weniger Interesse an MINT-Studiengängen und -berufen, was zu einem Ungleichgewicht in diesen bedeutsamen Fachbereichen führt. Zur Bewältigung dieser Herausforderung ist es von zentraler Bedeutung, die Ursachen dieser Differenzen zu ergründen und adäquate Strategien zu konzipieren, um die Förderung von Schülerinnen in den MINT-Disziplinen zu intensivieren und ihre Selbstwirksamkeitserwartungen zu konsolidieren. Eine Schlüsselrolle spielt dabei der Naturwissenschafts- bzw. Physikunterricht, der traditionell oft fachbezogen und rein theoretisch gestaltet wurde, ohne die Lebenswelt der Schülerinnen einzubeziehen (Archer et al., 2012; Hoffmann et al., 1998). Diese benötigen eine Umgebung, in der sie sich mit den Inhalten identifizieren können, damit ihr Interesse und ihre Motivation aufrechterhalten werden. Durch die Gestaltung eines geschlechtergerechten Unterrichts, der die Interessen und Bedürfnisse der Schülerinnen berücksichtigt, gelingt es, sie besser anzusprechen und ihre Selbstwirksamkeitserwartung zu stärken (Kosuch, 2010). Serious Games können hier vielversprechende Möglichkeiten bieten. Dabei handelt es sich um digitale Spiele, die speziell für Bildungs- und Lernzwecke entwickelt werden (Susi et al., 2007). Im Gegensatz zu herkömmlichen Unterrichtsmethoden erlauben Serious Games den Schülerinnen und Schülern, in einer interaktiven und ansprechenden Umgebung zu experimentieren und Probleme zu lösen. Durch die direkte Beteiligung am Spiel erhalten Lernende sofortiges Feedback und ein Gefühl der Kontrolle über ihre Lernfortschritte. Die Möglichkeit, verschiedene Lösungswege auszuprobieren, fördert sowohl die Interaktivität als auch das Flow-Erleben, einen Zustand vollständiger Vertiefung und intrinsischer Motivation. In diesem Zustand trainiert eine Person ihre Kompetenzen und erfährt gleichzeitig ein angemessenes Maß an Herausforderung (Csikszentmihalyi, 2014). Studien haben gezeigt, dass das Erleben von Flow in Serious Games dem Lernen und der Leistung förderlich sein kann (Cheng, Chen et al., 2015). Insbesondere Schülerinnen könnten von der Integration solcher Serious Games in den MINT-Unterricht profitieren. Durch deren Einsatz gelingt es, MINT-Inhalte auf spielerische und motivierende Weise zu vermitteln, was das Interesse und die Motivation steigern kann (Riopel et al., 2019). Darüber hinaus ermöglichen Serious Games, in einer sicheren Umgebung zu experimentieren und Fehler zu machen, ohne Angst vor negativen Konsequenzen zu haben (Bergey et al., 2015). Dies kann zu einem verbesserten Selbstvertrauen führen und das Risiko mindern, dass Schülerinnen aufgrund von Misserfolgen oder Vorurteilen demotiviert werden. Dieses Erfahrungsumfeld ermöglicht es Schülerinnen, ihre eigenen Fähigkeiten zu erkennen und ihr Selbstvertrauen aufzubauen, da sie unmittelbar erleben, dass ihr Handeln in der Spielwelt wirksam ist (Bandura, 1977; Garris et al., 2002). Dies kann zu einer Steigerung der Selbstwirksamkeitserwartung

führen (Lu & Lien, 2020; Peifer et al., 2020), was wiederum die Motivation und das Lernen positiv beeinflusst (J. Chen, 2007; Häußler & Hoffmann, 1995; C.-H. Wang et al., 2018).

Der Einsatz von Serious Games und die Förderung des Flow-Erlebens ist nicht nur für den MINT-Unterricht relevant, sondern könnte für andere Unterrichtsfächer von Interesse sein. Das Streben nach einem optimalen Lernerlebnis und einem Zustand des Flows sollte in allen Fächern sowie für sämtliche Schülerinnen und Schüler von Bedeutung sein. Jedoch zeigen nach wie vor Befragungen bei Lehrkräften, dass digitale Medien, trotz der drängenden Forderung (KMK, 2016; MKJS, 2020), bis heute im naturwissenschaftlichen Unterricht kaum Verwendung finden (Härtig et al., 2021). Ungeachtet der zunehmend verbesserten Gestaltung, die einen unkomplizierten Einsatz im Unterricht ermöglichen, greifen Lehrkräfte nach wie vor nur selten auf Serious Games zurück (Bos et al., 2014; Eickelmann et al., 2019). Dies ist einerseits auf die technische Ausstattung, andererseits auf das Wissen und die Einstellung der Lehrkräfte zurückzuführen (Bürger et al., 2021; Eickelmann et al., 2019; Wenzel, 2018). Diese sind laut den Ergebnissen besorgt, dass Serious Games Lerninhalte nicht ausreichend vermitteln können, nicht dem Lehrplan entsprechen und die Lernwirksamkeit nicht ausreichend empirisch belegt ist (Wastiau et al., 2009, S. 42–43). Nicht nur seitens der Lehrkräfte werden daher Methoden und Werkzeuge gefordert, um Serious Games wirksam in den Unterricht zu integrieren, sondern auch von Forschenden (Wouters et al., 2013). Angesichts der geforderten Digitalisierung ist die Entwicklung und Erforschung von begründeten fachdidaktischen Konzepten für die Unterrichtsentwicklung unabdingbar (KMK, 2016).

Diese Arbeit soll einen Beitrag hierzu leisten, indem ein Serious Game entwickelt und untersucht wurde, um speziell Schülerinnen anzusprechen und ihre Teilnahme im MINT-Bereich zu erhöhen. Durch das Serious Game sollen physikalische Konzepte der Elektrizitätslehre auf ansprechende und verständliche Weise vermittelt werden. Die Elektrizitätslehre ist ein grundlegendes Konzept in den Naturwissenschaften und spielt eine zentrale Rolle in unserem täglichen Leben. Jedoch kann sie aufgrund ihrer abstrakten Zusammenhänge und der Notwendigkeit, sowohl theoretische als auch praktische Anwendungen zu verstehen, eine Herausforderung für viele Lernende darstellen. Durch die spielerische Erfahrung und den interaktiven Charakter des Serious Games soll das Fachwissen der Lernenden verbessert werden. Aber auch die Rolle des Flow-Erlebens soll untersucht werden. Die Elemente des Spiels, die Flow-Erleben ermöglichen zeigen Übereinstimmungen mit den Elementen zur Stärkung der Selbstwirksamkeitserwartung, weshalb im Flow-Erleben eine Schlüsselrolle bei der Veränderung der Selbstwirksamkeitserwartung durch entsprechend gestaltete Serious Games vermutet wird. Des Weiteren soll überprüft werden, ob und wenn ja welche Rolle das Flow-Erleben bei der Wirkung der Prädiktoren Selbstkonzepte sowie Vorwissen auf die Änderung der Selbstwirksamkeitserwartung hat. Zusammenhänge zwischen Selbstkonzepten und Flow-Erleben wurden bislang nur im Zusammenhang mit sportlichen Tätigkeiten untersucht, nicht jedoch für Serious Games insbesondere im MINT-Kontext. Die Verbindung zwischen physikbezogenem sowie computerbezogenem Selbstkonzept und Flow-Erleben sowie die Wirkung des Flow-Erlebens auf das Lernen und die Selbstwirksamkeitserwartung zu erklären, ist von Bedeutung für die zukünftige Forschung, aber auch für die Entwicklung von Unterrichtsmaterial.

Um die Wirksamkeit des Serious Game zu evaluieren, wurde eine Interventionsstudie im Pre-Post-Design mit Kontrollgruppe (engl. *randomized controlled trial*) durchgeführt. In diesem Design werden Teilnehmende zufällig einer Experimentalgruppe oder einer Kontrollgruppe zugeordnet. Während die Experimentalgruppe das Serious Game spielt, erhält die Kontrollgruppe die gleichen Inhalte papierbasiert und mit Realversuchen, um die Auswirkungen einer Intervention zu messen. Mit der Anwendung dieses Designs könnten wir möglicherweise genauere Aussagen über die Wirksamkeit des

Serious Games machen und die Beziehungen zwischen Flow-Erleben, Fachwissen und Selbstwirksamkeitserwartung erforschen.

Die Struktur der Arbeit ist wie folgt: Nach einer Einführung in die theoretischen Grundlagen, die den Forschungsfragen zugrunde liegen, werden die verwendeten Versuchsmaterialien präsentiert. Hierbei wird nicht nur das Serious Game und die Materialien der Kontrollgruppe vorgestellt, sondern auch die Gestaltung der Fragebögen erläutert. Im Anschluss werden die verwendeten statistischen Methoden und das Untersuchungsdesign erläutert. Der Ergebnisteil stellt die Ergebnisse der Pilotierung sowie der Haupterhebung vor, die im Anschluss kritisch diskutiert und Limitationen aufgezeigt werden.

Die Ergebnisse sollen wichtige Erkenntnisse darüber liefern, ob Serious Games eingesetzt werden können, um Lernende mit geringen Selbstwirksamkeitserwartungen für MINT zu begeistern und ihre langfristige Teilnahme in diesen Bereichen zu fördern. Indem der Fokus auf das Flow-Erleben, das Fachwissen und die Selbstwirksamkeitserwartung gelegt wird, lassen sich innovative Ansätze entwickeln, um die Geschlechterkluft im MINT-Bereich zu verringern und eine von Inklusion und Vielfalt geprägte Zukunft für alle zu schaffen.

2 Serious Games in der (MINT)-Bildung

Serious Games dienen einem „expliziten und sorgfältig durchdachten Bildungszweck“ (Abt, 1971, S. 9). Es handelt sich um digitale Spiele, deren Nutzung mehr als Unterhaltung ist (Sawyer, 2007; Sawyer & Rejeski, 2002). Serious Games umfassen alle digitalen Spiele, die zum Training oder Lernen in Bereichen wie Gesundheit, Militär oder Schule eingesetzt werden können (Ratan & Ritterfeld, 2009). Dabei werden sie nicht nur zum Erreichen bestimmter Lernziele, sondern auch zur Beschreibung sowie Veränderung menschlichen Erlebens und Verhaltens genutzt (Brauner & Ziefle, 2022). Aufgrund ihrer Möglichkeit, Einfluss auf die Lernwirksamkeit zu nehmen, haben Serious Games im MINT-Unterricht immer mehr an Bedeutung gewonnen (Riopel et al., 2019). Die Lernwirksamkeit von Serious Games wird durch eine Kombination aus verschiedenen Theorien und Modellen erklärt. Die Motivation scheint dabei essenziell zu sein und kann erklären, warum Personen unterschiedlich intensiv mit Lerninhalten, aber auch mit Spielinhalten interagieren (Krath et al., 2021). Von besonderer Relevanz für die Erklärung der Motivation in Serious Games ist die *Flow-Theorie*, in der ein Zustand erhöhter Konzentration und Immersion beschrieben wird. Es wird vermutet, dass motivierte Schüler:innen eher bereit sind, sich länger mit Lernaufgaben auseinanderzusetzen und so Kompetenzen zu entwickeln (Fischer et al., 2021). Dieser Zusammenhang wird auch in Banduras Theorie von *Selbstwirksamkeitserwartung* beschrieben, dem Glauben an die eigenen Kompetenzen und Fähigkeiten (Bandura, 1977). Ein vermuteter Prädiktor dafür, wie motiviert sich Individuen mit den Aufgaben auseinandersetzen, ist das *Selbstkonzept*, d. h. die Eigenwahrnehmung einer Person in Bezug auf ihre Persönlichkeit inklusive ihrer Kompetenzen (Shavelson et al., 1976). Beispielsweise werden Personen, die um ihre guten Fähigkeiten in Computerspielen wissen, Aufgaben in diesem Zusammenhang möglicherweise mit höherer Motivation angehen. Aber auch Personen, die ihre Kompetenzen bezüglich spezifischer Inhalte des Serious Games als hoch einschätzen, könnten dieses mit einer größeren Motivation spielen. In den folgenden Abschnitten wird der Forschungsstand zu den Konstrukten *Flow-Erleben*, *Selbstwirksamkeitserwartung* und *Selbstkonzept* in Zusammenhang mit der Lernwirksamkeit von *Serious Games* dargelegt. Die Zusammenhänge werden abschließend in einem nomologischen Netzwerk aufgezeigt.

2.1 Serious Games – Spiele mit einem Bildungszweck

Die Idee, Lernvorgänge spielerisch zu gestalten, um die positive Wirkung der Motivation zu nutzen, wird bereits seit über 100 Jahren diskutiert (Groos, 1899). Systematisch untersucht wurde die Bedeutung von spielerischen Lernen zum ersten Mal von Huizinga (1938). Er beschreibt das Spiel als zweckfreie, freiwillige Handlung, deren Bedeutung in der Gegenwart liegt (ebd.). In dieser Definition ist bereits eine Ähnlichkeit mit dem Flow-Erleben (Csikszentmihalyi, 2014) zu erkennen (siehe auch Abschnitt 2.2). Mit dem Spielverhalten ist auch das Neugierverhalten verbunden, bei dem durch Exploration neue Erkenntnisse erlangt werden (Sachser, 2004). Damit können beide Verhaltensweisen als natürliche Voraussetzungen für das Lernen gesehen werden (vgl. Lengning, 2009). Spielen und Lernen hängen eng zusammen: In jedem Spiel findet Lernen statt, ohne dass der Wissenserwerb als solcher wahrgenommen wird. Dieses implizite Lernen gilt als ideal und erwünscht, während explizites Lernen den Spielfluss stören kann (Le et al., 2013). Andere Theoretiker argumentieren, dass Spielen seinen Zweck in sich selbst hat und ein Spiel mit einem Bildungsziel kein Spiel mehr ist (Caillois, 2001; Huizinga, 1938; Lazarus, 1883; Oerter, 2011). Jedoch wird das Spiel bei Kindern seit jeher für die Herausbildung von kognitiven oder emotionalen Kompetenzen genutzt (Ganguin, 2010). Es gibt zahlreiche Möglichkeiten, Spiele in die Bildung zu integrieren. Rollenspiele, Planspiele, Kartenspiele etc. können vielfältig im Unterricht eingesetzt werden und als Abwechslung beim (deklarativen)

Wissenserwerb dienen. Im Laufe der Zeit entstand die Idee, auch Computerspiele in die Bildung zu integrieren.¹ Die Ansätze und Definitionen in Bezug auf digitale Spiele mit einem Bildungszweck wurden in der Literatur vielfältig beschrieben und diskutiert (Breuer & Bente, 2010; Djaouti et al., 2011; Michael & Chen, 2006; Ratan & Ritterfeld, 2009; Wilkinson, 2016; Zyda, 2005). Dieser Arbeit soll die Definition zugrunde gelegt werden, wonach Serious Games digitale Spiele mit einem klaren Bildungsziel sind, die Spielenden Kompetenzen vermitteln, die auch über die Lernziele hinaus gehen (nach Marr, 2010, S. 18).

2.1.1 Lernwirksamkeit von Serious Games

Die Wirksamkeit von Serious Games als effektive Lehr- bzw. Lernwerkzeuge wurde bereits in unterschiedlichen Settings untersucht (Hoblitz, 2015; Lamb et al., 2018) und soll auch im Rahmen dieser Arbeit mit zwei gängigen Theorien erklärt werden: zum einen durch ihren Einfluss auf kognitive Prozesse - z. B. Lernen, Wissenserwerb und Problemlösen - zum anderen durch ihren Einfluss auf die Motivation (Wouters et al., 2013). Die Lern- und Motivationstheorien lassen sich zu einem Modell der Wirksamkeit von Serious Games zusammenführen. Die Mechanismen der Lernwirksamkeit von Serious Games werden durch unterschiedlichen Lerntheorien erklärt. Dabei wird die Mehrheit der Serious Games in der kognitivistischen Lerntheorie (Bruner, 1961) verortet und ein kleinerer Teil in der konstruktivistischen Lerntheorie (Dilling, 2022), dem Lernen durch Selbststeuerung und Einsicht (Kebritchi & Hirumi, 2008; Krath et al., 2021).² In der kognitivistischen Lerntheorie wird zwischen prozeduralem und deklarativem Wissen unterschieden. Bei Letzterem handelt es sich um Faktenwissen, das z. B. durch das Serious Game erworben wird, während sich prozedurales Wissen auf Kompetenzen bzw. Fertigkeiten bezieht, d. h., wie bspw. im Spiel gesteuert wird. Eine mögliche dritte Kategorie stellt das metakognitive Wissen (Wirth & Leutner, 2005) über die Steuerung und Wahl geeigneter Lernprozesse dar, die sich in Serious Games bspw. beim Lernen durch Identifikation und Integration neuer Informationen zeigen (Zumbach et al., 2020). Innerhalb der kognitivistischen Lerntheorie ist die sozial-kognitive-Lerntheorie mit ihrem Schlüsselkonzept, der Selbstwirksamkeitserwartung (Bandura, 1977; siehe Abschnitt 2.3) verortet. Die sozial-kognitive Lerntheorie beschreibt das Lernen als Prozess, bei dem Handlungen beobachtet, reflektiert und gegebenenfalls übernommen werden (Schunk, 2012).³

In dieser Arbeit liegt der Fokus auf den kognitiven Dimensionen des Lernens, insbesondere der Aneignung von Wissen z. B. aus Bildern oder Texten, aber auch auf komplexeren kognitiven Prozessen, wie dem Problemlösen, bei dem Wissen auf neue Situationen und Problemstellungen übertragen wird (Wouters et al., 2009).

Unterschiede in den Lernergebnissen können unter anderem auf motivationale Zustände zurückgeführt werden (Bai et al., 2020; Krath et al., 2021; Qian & Clark, 2016). So wird ein Einfluss der Lernmotivation (H. Heckhausen, 1969) vermutet, der „momentanen Bereitschaft eines Individuums,

¹ Der Einzug des spielerischen Lernens im Unterricht in die digitale Welt fand 1971 mit dem Spiel *The Oregon Trail*, einem textbasierten Adventure für den Geschichtsunterricht über das Pionierleben im 19. Jahrhundert, statt (Toppo, 2021).

² Eine vollständige Aufschlüsselung der zugrunde liegenden Lerntheorien in Serious Games, u. a. bei Krath et al. (2021), Kraiger et al. (1993) und Wouters et al. (2009).

³ Gelegentlich findet sich eine Beschreibung der sozial-kognitiven Lerntheorie als eine Kombination aus Sozialpsychologie und Kognitionspsychologie. In dieser Arbeit werden sozialpsychologische Komponenten wie die Interaktion innerhalb der Gruppe der Lernenden jedoch nicht Gegenstand der Arbeit sein und daher nicht diese Definition beinhalten.

seine sensorischen, kognitiven und motivationalen Fähigkeiten auf die Erreichung eines Lernziels zu richten und zu koordinieren“ (H. Heckhausen, 1969). Die Lernmotivation wird in der pädagogisch-psychologischen Forschung unter anderem auf die intrinsische Motivation zurückgeführt. In der Literatur finden sich verschiedene Definitionen der intrinsischen Motivation, wobei eine Aufschlüsselung der einzelnen Definitionen an dieser Stelle nicht stattfinden soll.⁴ Für diese Arbeit soll folgende Definition Anwendung finden: „Tätigkeiten, deren Anreiz im Vollzug selbst liegt“ (engl. „activity running by its own drive“, Woodworth, 1918, S. 70). Menschen sind bereit, Zeit und Energie in das Spiel zu investieren – nicht wegen äußerer Belohnungen, sondern weil das Spiel an sich lohnend ist (Malone, 1981). In vielen Studien konnte ein positiver Zusammenhang zwischen intrinsischer Motivation, Lernleistung und Lernwirksamkeit festgestellt werden (Eyupoglu & Nietfeld, 2019).

2.1.2 Empirische Untersuchung der Lernwirksamkeit in Serious Games

In der englischsprachigen Literatur existieren viele Essays, in denen die motivationale Wirkung von Serious Games vorhergesagt wird, die so ein Lernen *wie von selbst* ermöglichen sollen (Gee, 2003; Prensky, 2003). Jedoch blieb es zunächst bei dieser anekdotischen Evidenz, und die Forderungen nach Forschung wurden eindringlicher. Die ersten Forschungsergebnisse (Ke, 2009; Sitzmann, 2011; Vogel et al., 2006; Wouters et al., 2009), wonach Serious Games zumindest nicht lernwirksamer als normaler Unterricht sind, konnten auf unterschiedliche Lernmodelle und Forschungsmethoden zurückgeführt werden (O'Neil et al., 2005).

Durch ein besseres Verständnis der Wirkmechanismen, wie Interaktionsmöglichkeiten im Spiel, Motivations- und Lernmechanismen, konnte die Gestaltung der Serious Games zunehmend optimiert werden (Wouters et al., 2013). Die Entwicklungen in der Serious-Game--Forschung zeigen sich nicht nur in der jährlich zunehmenden Zahl an Veröffentlichungen zu diesem Thema, sondern auch die Ergebnisse der Metaanalysen geben Hinweise darauf, dass sich Serious Games als Alternative zum konventionellen Unterricht etablieren (Riopel et al., 2019; Zhonggen, 2019). Ebenso belegen die zunehmende Anzahl von Studien, die in der Metaanalyse von Connolly et al. (2012) berücksichtigt wurden (129 Studien von 2004 bis 2009), sowie die daran anknüpfende Metaanalyse von Boyle et al. (2016) (143 Studien zwischen 2009 und 2014), dass Serious Games mindestens ebenso effektiv wie herkömmlicher Unterricht in ihrer Lernwirksamkeit sind. In beiden Metaanalysen gab es genaue Kriterien in Bezug darauf, welche Studien eingeschlossen wurden. Ebenso wurde festgestellt, dass zwischen den beiden Analysen auch die Forschungskriterien strenger wurden. Insgesamt hatten in 12 der 143 Studien der Metaanalyse die Spiele ein naturwissenschaftliches Thema. Es gab Hinweise, dass die Nutzende von Serious Games mindestens genauso gut wie die Kontrollgruppe oder besser abschnitten, jedoch wurden keine genaueren Effektstärken berichtet (Boyle et al., 2016). Auch die Metaanalyse von Zhonggen (2019), in der 275 Publikationen der Jahre 2009 bis 2018 berücksichtigt wurden, zeigt einen Trend zu mehr Publikationen sowie positiven Ergebnissen durch den Einsatz von Serious Games (Zhonggen 2019). R. Mayer (2019) konnte in seinem Review bei 12 von 16 Studien aus dem naturwissenschaftlichen Bereich zeigen, dass in diesem Fall das Lernen mit Serious Games mit einer Effektstärke von $d = .70$ (Cohen, 1988)⁵ möglicherweise wirksamer ist als konventionelle Lehrmethoden (I. Mayer et al., 2014; R. Mayer, 2019). Klare Ergebnisse für oder gegen die Lernwirksamkeit von Serious Games zu finden, gestaltet sich als schwierig. Nach wie vor sind

⁴ findet sich aber z. B. in J. Heckhausen und Heckhausen (2006); Rheinberg et al. (1997).

⁵ Effektstärke Cohens d : .20 bis .40 schwacher Effekt, .50 bis .70 mittlerer Effekt, > .8 starker Effekt; η^2 : .01 kleiner Effekt, .06 mittlerer Effekt, .14 großer Effekt (Cohen, 1988).

Studiendesigns unterschiedlich, haben oft keine adäquate Vergleichsgruppe, messen keine Leistungen oder berichten keine Effektstärken (R. Mayer, 2019; Sitzmann, 2011).

Weiterhin fällt auf, dass den Studien uneinheitliche Untersuchungsdesigns und -ziele zugrunde liegen, sodass es schwierig bleibt, die Wirksamkeit von Serious Games festzustellen (Giessen, 2015). R. Mayer (2019) unterscheidet in Zusammenhang mit Serious Games drei Untersuchungsdesigns mit ihren zugehörigen Fragestellungen (siehe Tabelle 1). Hieraus ergeben sich die Heursforderungen und Überlegungen, die bei der Untersuchung der Wirksamkeit von Serious Games zu berücksichtigen sind und welches Studiendesign sich eignet.

Tabelle 1 Die drei Formen der Spielforschung, übersetzt aus R. Mayer (2019).

Untersuchung	Beschreibung	Forschungsfrage
Erforschen eines Mehrwerts	Vergleicht die Lernergebnisse einer Gruppe, die die Basisvariante eines Serious Games spielt, mit denen einer Gruppe, bei der das Serious Game um ein Feature erweitert wurde.	Wird durch die Ergänzung um das Feature mehr gelernt?
Erforschen der kognitiven Wirkung	Vergleicht die Verbesserung kognitiver Fähigkeiten einer Gruppe, die ein kommerzielles Spiel spielt, mit derjenigen einer Kontrollgruppe.	Wird durch das Spielen eines Spiels eine Fähigkeit verbessert?
Medien-vergleichende Forschung	Vergleicht die Lernwirksamkeit einer Experimentalgruppe, die ein Serious Games spielt, mit der bei einer Kontrollgruppe, der Lerninhalte konventionell vermittelt werden.	Lernen Individuen mehr mit einem Spiel oder in einem konventionellen Setting?

Eine Standard-Methode zur Ermittlung der Unterrichtseffektivität von Serious Games ist die Durchführung von Experimenten in randomisiert kontrollierten Experimenten im Pre-Post-Design, bei der die Kontrollgruppe ebenfalls angemessene Lernmaterialien, wie Aufgaben und Texte erhält (R. Mayer, 2019). Dieses Untersuchungsdesign, das auch in dieser Arbeit Anwendung findet, entspricht einer Medien-vergleichenden Forschung und kann zu dazu beitragen, die relative Wirksamkeit von Serious Games im Vergleich zu traditionelleren Lernmethoden zu bestimmen. (Boruch, 1997; Spybrook & Raudenbush, 2009). Darüber hinaus ist eine geeignete Kompetenz- bzw. Lernergebnismessung notwendig, denn die Forschung sollte sich nicht auf eine Selbstauskunft beschränken (R. Mayer, 2019). In einigen Metaanalysen wurde die geringe Anzahl an Untersuchungen als randomisiert kontrollierte Studien mit adäquaten Kontrollgruppen kritisiert (Riopel et al., 2019; Sitzmann, 2011). So fällt auf, dass in vielen Studien entweder keine Kontrollgruppe vorhanden ist oder die Kontrollgruppen unter anderen Bedingungen, z. B. mit reinen Lesetexten, lernen mussten, wodurch ein Vergleich mit einem interaktiven Serious Game erschwert wird (Sitzmann, 2011).

In ihrer Metaanalyse untersuchten Riopel et al. (2019) ausschließlich Serious Games aus den Bereichen Physik, Chemie, Lebenswissenschaften, Astronomie und Geophysik. Die Daten stammten aus 79 Medien-vergleichenden Studien. Dabei vermuteten die Wissenschaftler eine Überlegenheit der Serious Games gegenüber dem konventionellen Unterricht. Bei gleicher Unterrichtszeit schnitten Serious Games bezüglich deklarativem (Effektstärke $d = .34$) und prozeduralem Wissensgewinn und -erhalt ($d = .31$) besser ab. Diese Ergebnisse decken sich mit den Gesamteffektstärken voriger,

spezifisch naturwissenschaftlicher Metaanalysen von Wouters et al. (2013) und Sitzmann und Ely (2011). Des Weiteren wurde festgestellt, dass die Unterschiede zwischen der Experimental- und der Kontrollgruppe größer waren, wenn Letztere aktives ($d = -.19$) statt des passiven ($d = .38$) Lernmaterials erhielt (Sitzmann, 2011). Zwar stimmen diese Werte in ihrer Größe mit Ergebnissen aus vorigen Metaanalysen überein (Sitzmann & Ely, 2011), jedoch sind sie in ihrer Höhe nicht signifikant. Außerdem ermittelten Sitzmann und Ely (2011) in neueren Studien ab 2010 höhere Lernergebnisse im Vergleich zu älteren Untersuchungen, was sie mit der Verbesserung der Serious Games hinsichtlich Technologie, Design und pädagogischer sowie didaktischer Gestaltung im Laufe der Zeit in Verbindung bringen. Sie fanden außerdem heraus, dass das Fach bzw. die Disziplin dabei irrelevant ist. Es gibt zwar Hinweise für eine – nicht-signifikante – größere Effektstärke bei Serious Games mit physikalischen Inhalten, aber ihre Höhe wurde nicht angegeben (Riopel et al., 2019). Da es im Serious Game, das in dieser Arbeit zum Einsatz kommt, um Elektrizität und somit um ein physikalisches Thema geht, sind Studien in diesem Kontext von besonderem Interesse. Squire et al. (2004) untersuchten die Wirkung eines Serious Games zu den Grundlagen der Elektrostatik bei 96 Schüler:innen eines 8. Schuljahres. Dabei schnitt die Experimentalgruppe ($N = 61$) signifikant besser ab als die Kontrollgruppe ($N = 35$). Stege et al. (2012) untersuchten den Einsatz eines Serious Games zum Thema *Transistor* und konnten in ihrer Studie mit 187 Schüler:innen des 8. Schuljahres ebenfalls einen schwach signifikant höheren Lerneffekt ($p = .044$) für die Experimentalgruppe ($N = 97$) im Vergleich zur Kontrollgruppe ($N = 90$) ermitteln. Allerdings ist der Effekt mit einer Stärke von $\eta^2 = .022$ als klein zu bewerten (Cohen, 1988). Auch hinsichtlich der Motivation war das Ergebnis signifikant höher für die Experimentalgruppe ($p = .012$), mit einem kleinen Effekt von $\eta^2 = .034$. Darüber hinaus konnten Squire et al. (2004) ermitteln, dass Schüler beider Gruppen signifikant bessere Leistungsergebnisse als die Schülerinnen hatten ($p = .001$, $\eta^2 = .07$), die Schüler der Experimentalgruppe schnitten auch signifikant besser ab als die Schüler der Kontrollgruppe ($p = .037$, $\eta^2 = .05$). Die Interventionen von Veermans und Jaakkola (2019) mit einer Simulation elektrischer Schaltkreise enthielten Instruktionen und Arbeitsaufträge, die dazu dienen sollten, verbreitete Fehlvorstellungen über elektrische Schaltkreise aufzudecken und zu beseitigen. Die Instruktionen wurden zunehmend schwieriger und waren dem Spielerlebnis eines Serious Games nachempfunden (ebd.). Die Forschenden konnten ebenfalls gute Lernergebnisse für verschiedene Interventionsdesigns ermitteln ($d = .78$ bis $d = 1.24$, je nach Instruktion). Insgesamt konnten neuere Studien zeigen, dass gut konzipierte Serious Games das konzeptuelle sowie problemlösende Lernen fördern und darüber hinaus das Interesse am Fach wecken können (Riopel et al., 2019; Zhonggen, 2019)

2.1.3 Lernförderliche Elemente in Serious Games

Serious Games können zu effektiven Lernwerkzeugen werden, da sie, anders als konventionelle Unterrichtsettings, Lerninhalte sowohl erfahrungs- als auch problemorientiert, situiert, interaktiv und handlungszentriert vermitteln und sofortiges Feedback bieten (Boyle et al., 2011, S. 72). Darüber hinaus bereiten sie Lernende auf die Arbeitswelt des 21. Jahrhunderts vor (Holland et al., 2003). Durch eine Immersion in das Spielgeschehen (Cheng, She & Annetta, 2015; Mäyrä & Ermi, 2011) können die Nutzenden ihr Wissen und ihre Kompetenzen in einer sicheren und kontrollierten Umgebung üben und anwenden (Squire & Jenkins, 2003; Susi et al., 2007; Toh & Kirschner, 2020). Dabei ist zumindest bei Serious Games aus dem MINT-Bereich eine besonders realistische Darstellung des Spielinhaltes oder eine fesselnde Handlung nicht vorteilhaft. Eine fotorealistische und überladene Grafik lenkt die Spielenden von den Lernaufgaben ab und verursacht eine zusätzliche kognitive Belastung (Sweller, 2005). Dadurch kann die Lernwirksamkeit beeinträchtigt werden (Admiraal et al., 2011; Böhme et al., 2020; Clark et al., 2016; Ke & Abras, 2013; Kiili, 2005; Riopel et al., 2019; Virvou et al., 2005; Wouters

et al., 2013), aber auch Frustrationen sind möglich (Brom et al., 2011; Schrader & Bastiaens, 2012). Auch die narrative Rahmung ist essenziell (Dickey, 2006) und kann den Spielenden helfen, eine persönliche Relevanz im Spiel zu sehen (Klimmt, 2009). Es ist von Vorteil, wenn das Lernmaterial gut in die Handlung integriert wird und der Spielfluss nicht durch Lerngelegenheiten unterbrochen wird (Cheng & Annetta, 2012; Ullah et al., 2022). Dafür sollte die Erzählung einen passenden Kontext für das Lernmaterial liefern (Bellotti et al., 2012; Cheng et al., 2014) und den Spielenden die Möglichkeit bieten, die Kontrolle im Spiel zu übernehmen und nicht nur linear eine festgelegte Reihenfolge von Aufgaben abzuarbeiten (Ke & Abras, 2013; Kickmeier-Rust et al., 2006). Die Möglichkeit, den Ablauf oder das Tempo des Spiels zu kontrollieren, wurde bereits in verschiedenen Studien als lernwirksames Merkmal identifiziert (Gredler, 1996; Lawless & Brown, 1997; O'Neil et al., 2005; Vogel et al., 2006), was vermutlich auf das Flow-Erleben (siehe folgender Abschnitt) zurückzuführen ist. Die Kontrolle über die Handlung zu erhalten, wird als wesentlicher Aspekt des Spiels angesehen und trägt nachweislich zur Motivation der Spieler bei (Blumberg & Fisch, 2013). Mit der Kontrolle und Handlungsfähigkeit ist auch die Interaktivität verbunden. Sie bezeichnet das Phänomen, im Serious Game Eingriffs- und Steuerungsmöglichkeiten zu haben. Beispielsweise wenn Hebel und Schalter betätigt werden können, um ein Objekt zu untersuchen, oder wenn Räume gewechselt und Level wiederholt werden, um ein bestimmtes Ziel oder eine bestimmte Punktzahl zu erreichen. Dass sich die Art und Darstellung von Informationen an die Interessen und Anforderungen der Lernenden anpassen lassen ermöglicht zusammen mit dem Feedback individualisiertes Lernen (Ratan & Ritterfeld, 2009). Dadurch werden die Aufmerksamkeit und das Engagement der Lernenden aufrechterhalten (Moreno & Mayer, 2007; Strzebkowski & Kleeberg, 2002). Die geeignete Platzierung von Feedback und Hilfestellungen ist daher in Serious Games essenziell (Rahimi et al., 2022). Spielende meiden längere Ausführungen, bei denen deklaratives in prozedurales Wissen umgewandelt werden soll, und bevorzugen es, durch die Spielinteraktion selbst prozedurales Wissen zu generieren (Kerres et al., 2009). Die Phasen der Informationsgewinnung gehören zum expliziten Lernen und sollten möglichst kurz gehalten werden (Quandt et al., 2008). *Performance before Competence* bezeichnet das Einbauen von Tutorials und Hilfestellungen direkt im Spiel (Gee, 2007). Eine Möglichkeit der Unterbringung von Hilfestellung stellt die Kommunikation mit nicht spielbaren Charakteren, auf Englisch *non-player characters (NPCs)*, oder anderen Mitspielenden dar. Der Chat mit anderen Spielenden ist zu einem Erfolgsfaktor für Serious Games geworden (Arnab et al., 2015; Virvou et al., 2005). Selbst wenn ein Spiel nicht auf Zusammenarbeit ausgerichtet ist, tauschen sich die Spielenden oft über Taktiken und Lösungen aus (Dickey, 2011). Dass kollaboratives Lernen darüber hinaus die intrinsische Motivation von Schüler:innen erhöht, konnte in verschiedenen Studien nachgewiesen werden (Brom et al., 2011; C.-H. Chen, 2019; Liao et al., 2019).

2.1.4 Serious Games im Unterricht und ihre Auswirkungen auf die Lernmotivation

Die Einsatzsituation der Serious Games ist vermutlich ebenfalls von Bedeutung. Beispielsweise werden sie im Schulunterricht nicht freiwillig gespielt, wodurch sie vielleicht ihre (intrinsische) Motivationskraft verlieren (Wouters et al., 2013). Analog zu anderen Lehrmitteln hängt der Einfluss auf die Lernergebnisse auch bei Serious Games davon ab, wie sie konzipiert sind und in die Lernwelt der Schülerin bzw. des Schülers integriert werden (Lamb et al., 2018). So sollten Serious Games altersangemessen sein und neben den Interessen der Zielgruppe (Feierabend et al., 2020) auch die Anforderungen des Lehrplans berücksichtigen (Barzilai & Blau, 2014; Braghirolli et al., 2016). Daher schlagen einige Autor:innen zusätzlich zur geeigneten Integration in den Unterricht auch die Einbindung von Lehrkräften in die Konzeption von Serious Games vor (Ke & Abras, 2013) vor, um die

Aufmerksamkeit auf die Lerninhalte zu richten (Admiraal et al., 2011; Annetta et al., 2013; Lamb et al., 2018; Wouters et al., 2013).⁶

Insbesondere im Fach Physik kann der Einsatz von Serious Games die Chance bieten, den Schüler:innen abstrakte theoretische Konzepte auf motivierende Weise zugänglich zu machen. Durch die Möglichkeit des individualisierten Lernens und des Experimentierens können Schüler ein tiefes Verständnis erlangen. Speziell im Bereich der Elektrizitätslehre mit ihren abstrakten Modellen, anspruchsvoller Mathematik wie den Kirchhoffschen Gesetzen und unsichtbaren Vorgängen, die nur in ihrer Wirkung nachvollzogen werden können, sind Fehlvorstellungen möglich (siehe Unterabschnitt 4.1.1). Auch empfinden besonders Mädchen Hemmungen beim Experimentieren mit elektrischem Strom und haben Angst, Schaden anzurichten (M. Schwarzer & Tschauko, 2010). Wie dargelegt, können Serious Games helfen, die Hemmschwelle für das Experimentieren zu senken, das Wissen im angepassten Tempo zu vermitteln und so Fehlvorstellungen abzubauen.

Mithilfe von Serious Games lässt sich der Schulunterricht handlungszentriert und interaktiv gestalten. Die Menge der durch Entdeckung, Problemlösung und Schlussfolgerung gewonnenen Erkenntnisse (Lamb et al., 2018) wird durch kognitive, motivationale und emotionale Faktoren bestimmt. Durch das Potential, den Wissenserwerb an die kognitiven Bedürfnisse der Lernenden anzupassen und sie zu motivieren (Malone, 1981; O'Neil et al., 2005) was sich möglicherweise auch langfristig auf die Leistung auswirkt. Denn „motiviert Schüler wollen Fertigkeiten erwerben und sie anwenden, sie sind begeistert bei der Sache, haben Freude daran, Lernaufgaben zu lösen, und sind stolz, wenn sie Erfolg haben“ (Seel, 2000, S. 87). In diesem Kontext ist es wichtig, zwischen verschiedenen Arten von Motivation zu unterscheiden. Lernmotivation bezieht sich auf das allgemeine Interesse und die Bereitschaft eines Schülers, neue Fähigkeiten zu erlernen und anzuwenden. Intrinsische Motivation hingegen bezieht sich auf das Interesse und die Freude an der Aufgabe selbst, unabhängig von externen Belohnungen oder Strafen. Motivation im Allgemeinen kann sowohl intrinsisch als auch extrinsisch sein, wobei extrinsische Motivation durch äußere Anreize wie Noten oder Belohnungen getrieben wird. Dass sich Serious Games positiv auf die Motivation auswirken, wird in verschiedenen Theorien vermutet (Garris et al., 2002; Malone, 1981). Ein möglicher Einfluss auf die intrinsische Motivation wird dabei im Flow-Erleben vermutet (Csikszentmihalyi, 1975).

2.2 Die Rolle des Flow-Erlebens in Serious Games

Die intrinsische Motivation wird in vielen Studien zu Serious Games entweder mit der Selbstbestimmungstheorie⁷ (Deci & Ryan, 1985) oder der Flow-Theorie (Csikszentmihalyi, 1975) erklärt (Eyupoglu & Nietfeld, 2019). Die Flow-Theorie ist mit der Selbstbestimmungstheorie verwandt (Deci & Ryan, 2000) und weist deutliche Überschneidungen mit dem Konzept der intrinsischen Motivation auf

⁶ Lehrkräfte können zudem einen Einfluss auf das Flow-Erleben (Weimar, 2005) und die Selbstwirksamkeitserwartung (Bandura & Cervone, 1986; Klassen & Tze, 2014) haben, weshalb sie im Rahmen dieser Intervention weitestmöglich ausgeschlossen wurden.

⁷ Die Selbstbestimmungstheorie besagt, dass die intrinsische Motivation davon abhängt, inwieweit die drei psychologischen Bedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit befriedigt werden können (Deci & Ryan, 1985; Ryan et al., 2006) verwendeten die Selbstbestimmungstheorie, um die intrinsische Motivation in verschiedenen Spielkontexten zu untersuchen. Die intrinsische Motivation wurde dabei als für Computerspiele relevantes Kernelement der Motivation beschrieben und als Autonomie, Kompetenz und Immersion operationalisiert. Die Ergebnisse zeigten, dass die wahrgenommene Autonomie und Kompetenz im Spiel mit Spielfreude und Veränderungen des Wohlbefindens vor und nach dem Spiel verbunden waren. Die Wahrnehmung von Kompetenz und Autonomie hing möglicherweise mit der intuitiven Natur der Spielsteuerung und dem Gefühl der Immersion der Teilnehmenden zusammen(ebd).

(Eyupoglu & Nietfeld, 2019). In der Beschreibung der intrinsischen Motivation, dass die Betroffenen sich vollständig und wirksam in die Tätigkeit einbringen (Woodworth, 1918, S. 70), ist eine Übereinstimmung mit dem Flow-Erleben erkennbar, einer tätigkeitsbezogenen Motivation, die das reflexionsfreie Aufgehen in einer Aufgabe beschreibt. Csikszentmihalyi (1975) prägte den Begriff im Rahmen seiner Erforschung von intrinsisch motivierenden Faktoren, als er verschiedene Berufsgruppen über das Erleben ihrer Tätigkeiten befragte und feststellte, dass sich einzelne Beschreibungen der Erlebensweisen ähnelten (ebd.). Er beschrieb Flow als „holistic sensation that people feel when they act with total engagement“ (Csikszentmihalyi & Csikszentmihályi, 1992, S. 36).

Flow und Motivation

Flow und intrinsische Motivation

Motivation und Flow

Intrinsische Motivation und Flow

Csikszentmihalyi (2014) zufolge gibt es neun Komponenten, die Ursachen und Wirkungen von Flow-Erleben beschreiben, sie werden in der Tabelle 2 aufgelistet. Diese sind das Resultat von Faktorenanalysen der Ergebnisse von Messungen des Flow-Erlebens, bei denen festgestellt wurde, dass sich verschiedene Komponenten stärker ähnelten (Fournier et al., 2007; Hamari & Koivisto, 2014; Riva et al., 2017). Wie jede Komponente zum ganzheitlichen Flow-Erleben beiträgt, ist jedoch weniger klar. Auch konnte in den einzelnen Studien nicht nachgewiesen werden, dass sich Flow-Erleben ermitteln lässt, indem die Mittelwerte jeder Komponente summiert werden (Moneta, 2021). Vermutlich müssen beim Flow-Erleben nicht alle Komponenten zur selben Zeit auftreten. So können die Komponenten 1–3 als Ursachen oder Antezedenzen gesehen werden und die Komponenten 4–9 als Auswirkung (Sedig, 2007).⁸

Tabelle 2 *Komponenten des Flow-Erlebens ergänzt nach Csikszentmihalyi (2014).*

Nr.	Komponente	Beschreibung
1	Gleichgewicht zwischen Herausforderung und Fähigkeiten	Die Aufgaben entsprechen den Fähigkeiten bzw. Kompetenzen der Personen oder liegen minimal darüber.
2	klare Ziele	Den Personen ist zu jedem Zeitpunkt klar, was zu tun ist und durch welche Handlungen sie ein bestimmtes Ziel erreichen.
3	eindeutiges Feedback	Das tätigkeitsbezogene Feedback sollte klar, prägnant und zeitnah sein und den Personen helfen, Konzentration und Engagement für die Aufgabe aufrechtzuerhalten.
4	Verschmelzung von Handlung und Bewusstsein	Personen gehen in einer Tätigkeit auf, die Handlungen sind fließend, das Bewusstsein konzentriert.
5	Konzentrationsgefühl	Personen müssen sich vollkommen auf die Aufgabe konzentrieren.

⁸ Dieser Darstellung widersprechen Barthelmäs und Keller (2021) insofern, dass sie die Passung von Fähigkeit und Herausforderung als eine Voraussetzung für die Wahrnehmung von klaren Zielen und Feedback als Teil des Flow-Erlebens sehen und berufen sich auf die eigene Darstellung von Csikszentmihalyi (2014).

6	Gefühl der Kontrolle	Personen müssen das Gefühl haben, das Ergebnis der Aufgabe beeinflussen zu können.
7	Verlust des reflexiven Selbstbewusstseins	Personen sind so auf die Aufgabe konzentriert, dass sie sich ihrer selbst oder der Umgebung nicht bewusst sind.
8	kein Gefühl für die Zeit	Die Personen verlieren ihr Zeitgefühl, insbesondere scheint die Zeit schneller zu vergehen.
9	autotelische Erfahrung	Personen erledigen Aufgaben um ihrer selbst willen und nicht wegen externer Anreize. ⁹

Obleich sich die Aufteilung der Komponenten des Flow-Erlebens kaum verändert hat und diesbezüglich unter Forschenden weitgehend Einigkeit besteht (Engeser et al., 2021), finden sich in der Literatur unterschiedliche Ansätze, wie das Flow-Erleben operationalisiert werden soll (Abuhamdeh, 2021; Løvoll & Vittersø, 2014). Neben den Unterschieden bei der Identifizierung der Kernelemente des Flow-Erlebens gibt es auch Studien, in denen gar nicht von unterschiedlichen Komponenten ausgegangen wird und stattdessen nur eindimensionale Items zur Messung von Flow verwendet werden (L. H. Chen et al., 2010; Leung, 2020).

2.2.1 Flow-Erleben und Lernen

Allgemein wird angenommen, dass das Flow-Erleben die Leistungsfähigkeit einer Person positiv beeinflusst, da durch die damit verbundene erhöhte intrinsische Motivation das Vergnügen an einer Aktivität gefördert wird. Ein Flow-Erleben führt zu einer stärkeren und längeren Auseinandersetzung mit einem Lerngegenstand, denn durch das positive Gefühl beim Lernen sind Schüler:innen motivierter, aufmerksamer und konzentrierter (Csikszentmihalyi & LeFevre, 1989; Csikszentmihalyi & Schiefele, 1993; Kiili, 2005). Es wird vermutet, dass Flow-Erleben und die Leistung in Lernsituationen in einem reziproken Zusammenhang stehen, sodass auch gute Leistungen einen Einfluss auf das Flow-Erleben haben (Engeser & Rheinberg, 2008; Engeser et al., 2005). Zudem werden Aufgaben, bei deren Bearbeitung Flow erlebt wurde, auch in Zukunft mit hoher Motivation angegangen, was den Wissenserwerb indirekt beeinflusst (Csikszentmihalyi & Schiefele, 1993; Schiefele & Schreyer, 1994). Darüber hinaus wird vermutet, dass das Flow-Erleben auch direkt lernförderlich sein kann, da es große Übereinstimmungen zwischen den Flow-fördernden und den lernfördernden Bedingungen, z. B. die Komponenten 1–3 aus Tabelle 2, gibt (Csikszentmihalyi & Schiefele, 1993; Kiili, 2005).

Ob Flow erlebt wird, hängt vermutlich mit dem Vorwissen und der Anforderung des Lernstoffs zusammen. Ziele und Feedback müssen zudem eindeutig in den Lernprozess eingebunden sein. Lernenden muss klar sein, welcher Stoff in welcher Intensität beherrscht werden muss, und der Lernprozess muss durch Feedback eine eindeutige Handlungsstruktur aufweisen. Belege für einen möglichen positiven Zusammenhang zwischen Flow-Erleben und Lernwirksamkeit konnten Engeser und Rheinberg (2008) in einer Pre-Post-Untersuchung mit 197 Studierenden der Psychologie finden, die sich auf eine Statistik-Klausur vorbereiteten. Dabei wurden lediglich 4 % der Varianz der Leistung durch Flow-Erleben erklärt. Die Autor:innen schlossen daraus, dass Flow-Erleben ein Leistungsindikator ist und nicht nur Teil der Höchstleistung. Durch die Prädiktoren *Note in Mathematik*, *Alter* sowie *Vorwissen* konnten insgesamt 28 % der Varianz erklärt werden. Dabei hatten das Alter, die

⁹ Der bereits verwendete Begriff *Vergnügen* (Csikszentmihalyi & Csikszentmihályi, 1992) soll sich in dieser Komponente wiederfinden, da Tätigkeiten intrinsisch belohnend sein können, wenn sie Spaß machen (Abuhamdeh & Csikszentmihalyi, 2012).

Teilnehmenden waren zwischen 18 und 54 Jahre alt, und die Note in Mathematik den signifikantesten Einfluss. Bezüglich des Vorwissens zeigte sich, dass es nur einen marginal signifikanten Einfluss auf die Leistung hatten (Engeser & Rheinberg, 2008).

Die notwendigen Komponenten für ein Flow Erleben (siehe Tabelle 2, Komponente 1) untersuchten Hung et al. (2015) in einer Vergleichsstudie mit einem Lernspiel. Dabei bestand die Stichprobe aus 52 Schüler:innen des 2. Schuljahres. In zwei Gruppen wurde die Passung zwischen der Herausforderung und den Kompetenzen bzw. Fähigkeiten verglichen. Die Experimentalgruppe beschäftigte sich mit einem Spiel, dessen Herausforderung etwas höher war als die Kompetenzen bzw. Fähigkeiten, während bei der Kontrollgruppe die Passung optimal war. Die Forscher konnten nicht nur signifikant bessere Lernergebnisse, sondern auch signifikant höhere Werte beim Flow-Vorläufer ($\eta^2 = .12$) und beim Flow-Erleben ($\eta^2 = .15$) in der Experimentalgruppe feststellen.

Es gibt Hinweise darauf, dass das Flow-Erleben einen Einfluss auf den Wissenserwerb haben könnte, sowohl im Hochschulbereich als auch in der Schule (Barzilai & Blau, 2014). Beispielsweise untersuchte Franz (2020) in ihrer Arbeit die Mediatorrolle des Flow-Erlebens bei Schüler:innen der 7. bis 11. Schulstufen im Rahmen einer motivationsfördernden Intervention. Der Unterschied zwischen den Geschlechtern wurde nicht untersucht, sodass eine Forschungslücke erkennbar wird. Die Forscherin fand heraus, dass ein hohes Maß an Leistungsstreben sowie Ausdauer und Fleiß ein Prädiktor für ein hohes Flow-Erleben ist. Damit konnte sie die theoretischen Überlegungen unter anderem von Engeser und Vollmeyer (2005) bestätigen, dass die bei Schüler:innen verschiedenen ausgeprägten Facetten der Leistungsmotivation (Leistungsstreben, Ausdauer und Fleiß, Angst vor Misserfolg, aktivierende Prüfungsangst) signifikant das Flow-Erleben und damit das Wohlbefinden in der Schule bestimmen (Franz, 2020).

Vor allem Fächer wie Physik, die bei Schüler:innen im Allgemeinen, insbesondere aber bei Schülerinnen wenig beliebt sind (Welberg et al., 2021), könnten vom Flow-Erleben profitieren. Eine besondere Bedeutung für ein mögliches Flow-Erlebnis im Physikunterricht kommt den Experimenten zu (Barzel et al., 2012). Diese sorgen dafür, dass abstrakte Dinge erlebbar werden und der Bezug zur Lebenswelt der Lernenden hergestellt wird (Lechte, 2008). Berger (2018) untersuchte in einem rotierenden Untersuchungsdesign mit 407 Schüler:innen das Flow-Erleben beim Arbeiten mit realen und virtuellen Experimenten. Schülerinnen erlebten sowohl im Realversuch als auch im virtuellen Versuch hochsignifikant weniger Flow als Schüler, allerdings mit geringen Effektstärken (Realversuch $\eta^2 = .012$, virtueller Versuch $\eta^2 = .024$). Ebenfalls konnte gezeigt werden, dass im virtuellen Versuch hochsignifikant mehr Flow erlebt wurde als im Realversuch.

Die Rolle des Flow-Erlebens wurde nicht nur im konventionellen Unterricht untersucht (Egbert, 2004; Shernoff et al., 2014), sondern in den letzten Jahrzehnten zunehmend auch in Zusammenhang mit digitalen Spielen wie Serious Games (Schiefele & Roussakis, 2006). Dabei konnte in mehreren Studien zu digitalen Lerngelegenheiten der direkte Zusammenhang zwischen Flow-Erlebnis und Zufriedenheit der Schüler:innen bzw. Studierenden belegt werden (Rossin et al., 2009). Studien mit Serious Games kamen zu dem Ergebnis, dass sich Flow-Erfahrungen positiv auf das Engagement, die Motivation und die Lernergebnisse auswirken (Hoblitz, 2015; Kiili, 2005). Im Folgenden soll der Fokus ausschließlich auf Studien mit einem Prä-Post-Design mit Kontrollgruppe liegen, die in ihrer Art und Durchführung Ähnlichkeiten mit der vorliegenden Arbeit aufweisen.¹⁰

Die Studie von Hamari et al. (2016) umfasste eine Stichprobe von $N = 173$. Dabei wurden zwei Lernspiele eingesetzt und das Flow-Erleben über die Flow-Komponente 1 (siehe Tabelle 2) als

¹⁰ Eine Übersicht auch über Studien, die Flow im Zusammenhang mit Lernwirksamkeit in Spielen allgemein untersucht, findet sich z. B. bei Perttula et al. (2017).

Gleichgewicht zwischen Herausforderung und Fähigkeiten operationalisiert. Des Weiteren wurden das Engagement und die Immersion erhoben. Engagement wird als Zustand beschrieben, in dem Schüler:innen motiviert an einer Tätigkeit teilnehmen, was zu effektivem Lernen führt (ebd.). Immersion bezieht sich auf das Gefühl, vollständig in einer Aktivität oder Erfahrung aufgehen zu können. Bei dieser zeigte sich kein Zusammenhang mit der Lernwirksamkeit, während das Engagement die Lernwirksamkeit positiv beeinflusste. Die Herausforderung war ein stark signifikanter Prädiktor für die Lernwirksamkeit. Bezüglich der Gestaltung von Lernspielen wird empfohlen, dass bei der Herausforderung des Spiels die wachsenden Fähigkeiten und der Lernfortschritt berücksichtigt werden sollten, um das kontinuierliche Lernen zu unterstützen (ebd.). Ist das Spieldesign nicht auf das Können der Spielenden zugeschnitten, kann dies zu Frustration und Demotivation führen, infolgedessen sich die Lernwirksamkeit verringert (ebd.). Daher sollten Spieldesigner und Pädagogen darauf achten, ein Gleichgewicht zwischen Herausforderung und Kompetenz zu schaffen, das einen gesunden Flow-Zustand fördert, und es den Spielenden ermöglichen, ihr Gelerntes in Situationen der realen Welt anzuwenden. (ebd.).

Auch im MINT-Kontext wurde das Flow-Erleben bei Serious Games untersucht. Tsai et al. (2016) konnten in ihrer Studie mit einem Serious Game zum Thema *Elektromagneten* Zusammenhänge mit dem Vorwissen und dem Flow-Erleben aufdecken. Zwischen den Lernenden mit unterschiedlichen Vorkenntnissen gab es signifikante Unterschiede hinsichtlich der Flow-Erfahrung und der kognitiven Belastung. So neigte die verständnischwache Gruppe dazu, die Nachrichten wiederholt zu lesen, ohne spezifisches weiteres Verhalten zur Lösung der Aufgabe zu zeigen. Das deutet auf einen Mangel an metakognitiven Kontrollstrategien in der spielbasierten Lernumgebung hin. Dieses Ergebnis bestätigt die Berichte früherer Studien, dass es signifikante Korrelationen zwischen Flow, Lernerfolgen und kognitiven Strategien beim Lernen mit digitalen Medien gibt (Admiraal et al., 2011; Tsai et al., 2016). Lernstrategien haben möglicherweise Einfluss auf die intrinsische Motivation (Schiefele & Schreyer, 1994).

Auch Rachmatullah et al. (2021) konnten in ihrer Studie mit einem Serious Game zum Thema *Genetik* bei Mittelschüler:innen große Lernfortschritte durch das Flow-Erleben zeigen. Sie ermittelten außerdem, dass Vorwissen keinen signifikanten Einfluss auf das Flow-Erlebnis der Schüler:innen hatte, aber signifikant mit Frustrationsgefühlen verbunden war. Gee (2007) sieht die angenehme Frustration („pleasantly frustrating“, S.36) als einen positiven Aspekt von Spielen. Ein gewisses Maß an wiederholter Frustration kann die Schüler:innen auch darauf vorbereiten, dieses Gefühl zu tolerieren, das in der Bildung und im Leben im Allgemeinen häufig auftritt (Shute et al., 2015). Serious Games haben das Potential, den Schüler:innen unabhängig von Vorwissen und Vorerfahrung ein Flow-Erlebnis zu ermöglichen, die intrinsische Motivation zu stärken und lernwirksam zu sein, indem sie die Merkmale berücksichtigen, die mit dem Flow-Erleben in Zusammenhang stehen (Kiili, 2005; Nacke & Lindley, 2009).

2.2.2 Merkmale in Serious Games für ein Flow-Erleben

Die Merkmale des Flow-Erlebens lassen sich mit den Aspekten von Serious Games in Verbindung bringen. Serious Games haben klare Spielregeln sowie festgelegte Ziele (zweite Komponente), enthalten eine Feedback-Funktion (dritte Komponente) und der Schwierigkeitsgrad steigert sich mit den implizit erlernten Spiel-Fähigkeiten (erste Komponente; J. Chen, 2007). Durch die Immersion in Serious Games können die Spielenden ihr Zeitgefühl verlieren (achte Komponente, Nordin et al., 2013). Durch die Immersion in Serious Games sind die Spielenden so auf die Aufgabe konzentriert, dass sie sich ihrer selbst oder der Umgebung nicht bewusst sind (7. Komponente, Michailidis et al., 2018). Die Möglichkeit, den Verlauf des Serious Games zu kontrollieren, wird dabei als bedeutendster Einfluss auf

ein positives Spielerlebnis genannt (Kiili et al., 2012). Ist die Aufgabe auf das Können des Einzelnen abgestimmt, führt dies zu einem Gefühl der Kontrolle und Beherrschung (sechste Komponente; Alexiou & Schippers, 2018; Hamari et al., 2016). Durch Feedback wird den Spielenden Rückmeldung über ihren Fortschritt und ihre bisherigen Leistungen gegeben, was einen positiven Einfluss auf die intrinsische Motivation haben soll (dritte Komponente; Burgers et al., 2015).

Die Motivation könnte nicht nur einen Einfluss darauf haben, ob und wie intensiv sich jemand mit Lerninhalten auseinandersetzt, sondern auch auf das Gefühl, wirksam zu sein. Spielende sind auf verhaltensbezogener, kognitiver und emotionaler Ebene engagiert, was zum Erreichen von Lernzielen beiträgt, selbst in Bereichen, in denen die wahrgenommene Kompetenz eher gering ist (vierte Komponente, Plass et al., 2015). Spielende erleben, dass sie sich im Verlauf des Spiels verbessern. Sie können z. B. Rätsel schneller lösen, da sie Inhalte besser verstehen, wodurch auch ihr Glaube daran steigt, zukünftige Rätsel besser lösen zu können: Die Selbstwirksamkeitserwartung (Bandura, 1977) dieser Spielenden wird verbessert (Plass et al., 2015).

2.3 Wirkungen der Serious Games auf die Selbstwirksamkeitserwartung

Die sozialkognitive Perspektive der Selbstwirksamkeitserwartung (Bandura, 1977) beschreibt die eigene Überzeugung, wie gut eine Handlung organisiert und ausgeführt werden kann (Schunk & Mullen, 2012). Die Selbstwirksamkeitserwartung ist kontextspezifisch, d. h., sie bezieht sich auf eine spezifische Situation in einem bestimmten Bereich und kann nicht global wahrgenommen werden (Bandura, 1977; Zulkosky, 2009). Beispielsweise bezieht sich die schulische Selbstwirksamkeitserwartung auf die Überzeugung von Einzelpersonen, bestimmte schulische Aufgaben in spezifischen Bereichen erfolgreich erfüllen zu können (Schunk, 1991). Sie gilt als bedeutender Prädiktor für den Erfolg im schulischen und akademischen Bereich. Mit ihrer Hilfe lassen sich Lernerfolge im Allgemeinen (Zimmerman, 2000), und auf MINT-Fächer angewandt im Besonderen, voraussagen (Eddy & Brownell, 2016). Es konnte ermittelt werden, dass die schulische Selbstwirksamkeitserwartung mit der Leistung steigt (Bandura & Locke, 2003). Damit beeinflusst sie nachhaltig sowohl den Unterricht als auch die anschließende Leistungsentwicklung (Bandura, 1977; Nissen, 2019).

Personen mit einer hohen Selbstwirksamkeitserwartung haben auch eine hohe Frustrationstoleranz und sind gegenüber Misserfolgen vorübergehend widerstandsfähig (Zimmerman, 2000). Im Allgemeinen bemühen sich Menschen mit einer hohen Selbstwirksamkeitserwartung eher, eine Aufgabe zu erledigen, und zeigen dabei eine größere Ausdauer als Menschen mit einer geringen Selbstwirksamkeitserwartung (Schunk, 1990). Daher ist es von besonderem Interesse, die Selbstwirksamkeitserwartung, speziell in den unbeliebten MINT-Fächern, zu steigern. Nach Bandura (1977) beeinflussen insbesondere vier Quellen die Selbstwirksamkeitserwartung (siehe Tabelle 3).

Tabelle 3 Die vier Quellen der Selbstwirksamkeitserwartung nach Bandura (1977).

Nr.	Kurzname	Beschreibung
1	eigene Erfolgserlebnisse	Erfolgs-erlebnisse bei der Bewältigung schwieriger Aufgaben steigern den Glauben an die eigene Kompetenz in einem fachspezifischen Bereich (Bandura, 1977).
2	stellvertretende Erfahrung	Auch wenn Selbstwirksamkeitsüberzeugungen in der Regel stärker von der direkten Erfahrung beeinflusst werden (Zimmerman, 1995), kann in bestimmten Situationen auch der soziale Vergleich von wesentlicher Bedeutung sein. Die Erfahrung von Menschen mit ähnlichen Kompetenzen kann z. B. dann einen positiven Einfluss auf die Selbstwirksamkeitserwartung haben, wenn es an eigenen Vorerfahrungen mit der zu bewältigenden Aufgabe mangelt (Bandura, 1977). Schunk et al. (1987) konnten zeigen, dass das Gefühl der Wirksamkeit bezüglich einer Aufgabe bei Kindern, die in der Vergangenheit bei der Aufgabe erfolglos waren, durch stellvertretende Erfahrungen erhöht werden kann. Darüber hinaus ist die verbale Überzeugung als Quelle der Selbstwirksamkeitserwartung von hoher Bedeutung (Bandura, 1977). Unterstützende Botschaften können dazu dienen, die Bemühungen und das Selbstvertrauen zu stärken, insbesondere wenn sie von Bedingungen und Anweisungen begleitet werden, die zum Erfolg beitragen (Evans, 1989). Für eine bessere Wirksamkeit sollte das Feedback so formuliert sein, dass das Selbstvertrauen der Individuen gestärkt wird, vor allem wenn sich ihr Selbstvertrauen noch entwickelt (Schunk, 1984). Jedoch geben Usher und Pajares (2008) zu bedenken, dass bei Personen mit einer pessimistischen Einstellung durch verbales Feedback auch die Selbstwirksamkeitserwartung geschwächt werden kann. Solche Personen sehen Fehler als Zeichen ihrer Unfähigkeit (Seligman et al., 1990). Daher ist es insbesondere in den Jahren des Heranwachsens entscheidend, wie die Botschaften an die Jugendlichen formuliert werden (Bandura, 1977).
3	verbale Überzeugung	
4	emotionale Erregung	Durch positive Botschaften werden die Motivation und die anschließende Leistung gesteigert, und dieser wechselseitige Prozess erhöht das Wohlbefinden (Usher & Pajares, 2008). Diese Wahrnehmung der eigenen Gefühle ist die letzte von Bandura (1977) ermittelte Quelle der Selbstwirksamkeitserwartung. Eine Angstentwicklung in bestimmten Situationen führt zu einer Herabsetzung der Kompetenzerwartung (ebd.).

2.3.1 Die vier Quellen der Selbstwirksamkeitserwartung in Serious Games als Forschungsgegenstand

Die Gestaltung von Serious Games kann die vier Quellen der Selbstwirksamkeitserwartung (siehe Tabelle 3) berücksichtigen. Die Erfahrung, wirksam zu sein durch eigene Erfolgserlebnisse, lässt sich beim Serious Game durch das Erreichen von Zwischenzielen, das Abschließen von Aufgaben und das Spiel selbst realisieren (Urbani et al., 2002). Das Erreichen anspruchsvoller Ziele steht vermutlich mit einer gesteigerten Selbstwirksamkeitserwartung in Verbindung (Lieberman, 2006). Die Erledigung virtueller Aufgaben und die Beherrschung von Inhalten in der virtuellen Umgebung können zu einer

Steigerung der Selbstwirksamkeit in der realen Welt führen (Hardin et al., 2014). Die stellvertretende Erfahrung kann durch Avatare ermöglicht werden (Alserri et al., 2018). Bereits die Erstellung des Avatars bietet eine Chance zur Erhöhung der Selbstwirksamkeitserwartung, da die Anpassung der virtuellen Präsenz an das Selbst zu einer stärkeren Identifikation mit der Spielfigur führen soll (Alserri et al., 2018; Mitra & Golz, 2016; Yee & Bailenson, 2007). Die Quelle der verbalen Überzeugung stellen in digitalen Spielen die Belohnungssysteme dar, die das Engagement der Spielenden fördern sollen (Reeves & Read, 2009). Elemente wie Fortschrittsbalken (Gronier & Baudet, 2019), Badges und Punkte (J. C. Yang et al., 2016) und Ingame-Gegenständen, die direkt mit Lernergebnissen verknüpft werden, motivieren nicht nur, sondern verbessern die Lernwirksamkeit (Kiili, 2005; Sweetser & Wyeth, 2005) und tragen möglicherweise mit ihrem direkten Feedback bezüglich der Leistung auch zu einer höheren Selbstwirksamkeitserwartung bei (Gnauk et al., 2012).

Die Wirkung der verbalen Überzeugung kann zusätzlich erhöht werden, wenn das Serious Game in den Unterricht eingebunden und durch zusätzliches Feedback der Lehrkraft unterstützt wird (Alserri et al., 2018; Annetta et al., 2013). Die emotionale Erregung, die Schüler:innen beim Spielen von Serious Games erleben, kann als vierte Quelle vermutlich zu einer höheren akademischen Selbstwirksamkeitserwartung führen (Potosky, 2002). Individuen bemühen sich zu lernen, wenn sie sich als kompetent genug zum Lösen der Aufgabe halten und glauben, dass ihre Bemühungen zum Erfolg führen werden. Deshalb sollten Serious Games möglichst viele Elemente enthalten, die diese Wirkung aufrechterhalten und verstärken (Lee, 2000) oder zumindest eine erhebliche Abnahme der Selbstwirksamkeitserwartung vermeiden.

Bei Frust und anderen negativen Emotionen, die das Erlebnis dominieren, verliert das Individuum den Spaß am Spiel (Klimmt & Hartmann, 2006). Die Selbstwirksamkeitserwartung stellt daher für manche Forschende eine Erklärung der Motivation in Serious Games dar (I. Mayer et al., 2014; Wentzel & Miele, 2009). Es besteht die Möglichkeit, dass sich Menschen auf bestimmte Spiele nur dann (freiwillig) einlassen, wenn sie von einem Erfolg ausgehen (Vorderer et al., 2004). Die Motivation für Spiele würde dann von der Selbstwirksamkeitserwartung beeinflusst werden. Damit wäre auch das Geschlecht ein entscheidender Faktor, da bezüglich der Computerspiele von einer höheren Motivation seitens der Jungen ausgegangen werden kann (Janneck et al., 2012). In Zusammenhang mit Banduras Quellen der Selbstwirksamkeitserwartung konnte ermittelt werden, dass Jungen stärker von den Erfolgserlebnissen bei der Bewältigung von Aufgaben profitieren, während bei Mädchen die stellvertretenden Erfahrungen und die Wahrnehmung der eigenen Gefühle die Selbstwirksamkeitserwartung stärker beeinflussen (Usher & Pajares, 2008; Webb-Williams, 2018).

2.3.2 Empirische Befunde für die Veränderung der Selbstwirksamkeitserwartung durch Serious Games

Hinweise zu Einflüssen auf die Selbstwirksamkeitserwartung durch digitale Spielumgebungen lieferte die Studie von Ketelhut (2007), die Daten von 96 Schüler:innen des 7. Schuljahres beim spielerischen Einstieg in die wissenschaftliche Tätigkeit erhob. Die Forscherin konnte zeigen, dass sich die Selbstwirksamkeitserwartung in Bezug auf das Lernen durch die Nutzung einer virtuellen Spielumgebung insgesamt verbessert. Darüber hinaus wurde beobachtet, dass sich Schüler:innen mit einer geringen Selbstwirksamkeitserwartung nach wiederholter Nutzung der Spielumgebung mit den Lerninhalten gleich intensiv auseinandersetzen wie Schüler:innen mit einer hohen Selbstwirksamkeitserwartung (Ketelhut, 2007). Sie erklärt dies damit, dass die Schüler:innen im Pre-Test ihre schulische Selbstwirksamkeitserwartung geringer einstufen, da sie durch den Schulunterricht geprägt waren. Als weitere Möglichkeit zieht sie die eigenen motivierenden Wirkungen des Spiels in Betracht, die zu einer stärkeren Auseinandersetzung mit den Lerninhalten führte. Sie

konnte zwischen den Geschlechtern keine Unterschiede in der Selbstwirksamkeitserwartung feststellen. Jedoch fanden in einer späteren Untersuchung desselben Datensatzes B. C. Nelson und Ketelhut (2008) heraus, dass Schülerinnen häufiger die Hilfestellungen des Spiels abrufen, was einen relevanten Hinweis für die Gestaltung von Spielen gibt. Einen weiteren Beleg für die ausgleichenden Wirkungen der Merkmale von Serious Games ergab die Studie von Orvis et al. (2008), in der die computerspielbezogene Selbstwirksamkeitserwartung untersucht wurde. Sie stellten fest, dass erfahrene Computerspieler:innen anfangs von einer höheren aufgabenbezogenen Selbstwirksamkeit ausgehen und sich höhere persönliche Leistungsziele für die Aufgaben setzen, während unerfahrene Spieler:innen anfangs glauben, nur geringere Leistungsziele erreichen zu können (ebd.). Nach dem Kontakt mit dem Spiel durch mehrere Übungsversuche waren jedoch am Ende der Intervention die beobachteten Unterschiede zwischen Anfänger:innen und erfahrenen Spieler:innen nicht signifikant (ebd.). Hieraus kann geschlossen werden, dass auch Individuen, die zuvor wenig Kontakt zu Computerspielen hatten, vom Serious Game der vorliegenden Studie profitieren können.

Auch Hoblitz (2015) konnte in ihrer Feldstudie mit einem Serious Game zu physikalisch-technischen Inhalten mit $N = 176$ Schüler:innen des 9. Schuljahres hinsichtlich der Geschlechterunterschiede zeigen, dass Jungen häufiger Games spielen und dadurch eine höhere Selbstwirksamkeitserwartung beim Spielen haben. Darüber hinaus finden sie Physik interessanter. Mädchen haben dafür eine positivere Einstellung gegenüber Serious Games, was als ausgleichender Effekt angenommen wurde, da es kaum Unterschiede in den Ergebnissen des Fachwissenstests gab (ebd.). Den Zusammenhang zwischen der computerbezogenen Selbstwirksamkeitserwartung und der Spielerfahrung sowie dem Lernerfolg untersuchten Rosenthal et al. (2022) in einer Studie mit 241 Studierenden, die ein naturwissenschaftliches Serious Game spielten. Die Forscher fanden heraus, dass der Spielfortschritt mit der computerbezogenen Selbstwirksamkeitserwartung und dem Lernzuwachs in hoch signifikantem Zusammenhang steht. Dabei fungiert der Spielfortschritt als Mediator zwischen der computerbezogenen Selbstwirksamkeitserwartung und dem Lernzuwachs. Darüber hinaus ermittelten sie einen hoch signifikanten Zusammenhang zwischen der Erfahrung mit Computerspielen und der Selbstwirksamkeitserwartung. Zwar korreliert die Selbstwirksamkeitserwartung nicht mit dem Lernzuwachs, aber stattdessen der Lernzuwachs mit der ebenfalls erhobenen *gewöhnlichen naturwissenschaftlichen Intelligenz* ($r = .15$)¹¹, die in diesem Fall möglicherweise einen höheren Einfluss auf das Lernen hat als die Selbstwirksamkeitserwartung. Insgesamt kann daraus geschlossen werden, dass erfahrene Computer-Spielende vom Spiel profitieren konnten, da sie zuversichtlich waren, dieses lösen zu können und über den Spielverlauf zu lernen. Dafür lernten Spielende mit guten naturwissenschaftlichen Kompetenzen oder Fähigkeiten direkt aus dem Spiel (Rosenthal et al., 2022). Eine bedeutende Erkenntnis hinsichtlich des Vorwissens konnten (J. Wang et al., 2022) in einer neueren Studie mit einem Spiel zur newtonschen Physik gewinnen. Sie kamen zu dem Ergebnis, dass weder eine geringe Selbstwirksamkeitserwartung noch Vorwissen für den Lernerfolg maßgeblich sind. Spielende mit hoher Selbstwirksamkeitserwartung wandten in der Untersuchung erfolgreiche Spielstrategien an, während sich Spielende mit viel Vorwissen stärker auf die Lerninhalte konzentrierten.

In einem Review, das die Erkenntnisse der Selbstwirksamkeits-Forschung über einen Zeitraum von 20 Jahren umfasst, wurde die Selbstwirksamkeitserwartung als stärkster Prädiktor für die Lernmotivation ermittelt (Colquitt et al., 2000). Auch ein möglicher Zusammenhang zwischen der

¹¹ Interpretation von r nach Cohen (1988) $|r| = .10$ schwache Korrelation, $|r| = .30$ moderate Korrelation, $|r| > .50$ starke Korrelation.

naturwissenschaftsbezogenen sowie der technologiebezogenen Selbstwirksamkeitserwartung und dem Flow-Erleben bei herausfordernden Serious Games wurde herausgefunden (Lu & Lien, 2020).

2.3.3 Selbstwirksamkeitserwartung und Flow-Erleben

Die Selbstwirksamkeitserwartung, die eine signifikante Beziehung zum Flow-Erleben hat, könnte eine Schlüsselkomponente sein, um die Motivation in Spielen zu gewährleisten. Durch diese Motivation könnten Serious Games die Fähigkeit haben, Lernumgebungen zu verbessern (Pavlas, 2010; Whitson & Consoli, 2009). Selbstwirksamkeitserwartungen werden mit zentralen Motivationskonstrukten wie Ursachenzuschreibung, Optimismus oder Leistungszielorientierung in Verbindung gebracht (Usher & Pajares, 2008), die den Umfang der für eine Tätigkeit aufgewendeten Anstrengungen und die diesbezügliche Ausdauer beeinflussen (Bandura, 1993). Es gibt einen erkennbaren Zusammenhang zwischen der Selbstwirksamkeitserwartung und den Komponenten des Flow-Erlebens. Diese Beziehung kann in beide Richtungen wirken: Die Komponenten des Flow-Erlebens können einen positiven Einfluss auf die Selbstwirksamkeitserwartung haben und umgekehrt kann eine hohe Selbstwirksamkeitserwartung das Flow-Erleben positiv beeinflussen

(Rodríguez-Sánchez et al., 2011). Die Selbstwirksamkeitserwartung steht in Zusammenhang mit Voraussetzungen des Flow-Erlebens (siehe Abschnitt 2.2, Tabelle 2), wie z. B. klaren Zielen (zweite Komponente, s. Tabelle 2: Pintrich & Zusho, 2007). Personen mit hoher Selbstwirksamkeitserwartung setzen sich höhere Ziele, die mit Anstrengung und Ausdauer verbunden sind und nutzen ihre Selbstwirksamkeitserwartung als Ressource der Motivation (R. Schwarzer, 1998). Durch Feedback, eine der zentralen Voraussetzungen des Flow-Erlebens (dritte Komponente, s. Tabelle 2), kann die Motivation aufrechterhalten werden, höhere Ziele zu erreichen und Lernenden signalisiert werden, dass ihr Erfolg auf eigene Arbeit zurückzuführen ist (Kiili, 2005; Landhäußer & Keller, 2012; Locke & Latham, 1990). Dies hat einen möglichen Einfluss auf die Selbstwirksamkeitserwartung (Bandura & Schunk, 1981; Schunk et al., 1987). Durch den steigenden Schwierigkeitsgrad wird die Passung zwischen Herausforderung und Können (erste Komponente, s. Tabelle 2) gewährleistet und Spieler erleben Flow (Greenfield, 1994). Durch die Bewältigungserfahrung steigert sich die Selbstwirksamkeitserwartung, was eine motivierende Voraussetzung ist, um weiterzuspielen (Starks, 2014).

In ihrer Meta-Analyse mit 55 Vergleichsstudien konnte Sitzmann (2011) eine mögliche Zunahme der Selbstwirksamkeitserwartung ($d = .52$) sowie des prozeduralen ($d = .28$) und deklarativen Wissens ($d = .37$) – auch langfristig – durch das Spielen von Serious Games ermitteln (Sitzmann, 2011). Die Forscherin vermutet zwar, dass das Flow-Erleben das bedeutendste Element der motivierenden Wirkung von Serious Games ist, jedoch untersuchte sie die Daten der Studien nicht in diese Richtung. Der Zusammenhang von Selbstwirksamkeitserwartung und Flow-Erleben in Serious Games wurde von Pavlas et al. (2010) in einer Studie mit 120 Studierenden am Beispiel eines Biologie-Spiels untersucht. In einem ersten Modell war Flow ein signifikanter partieller Mediator für die Beziehung zwischen Selbstwirksamkeitserwartung und deklarativem Wissen ($R^2 = 29.06\%$). In einem zweiten Modell zeigte sich, dass das Flow-Erleben die Auswirkungen der Selbstwirksamkeitserwartung auf die intrinsische Motivation erklärt ($R^2 = 26.01\%$). Die Autor:innen ziehen den Schluss, dass die Selbstwirksamkeitserwartung eine wichtige Rolle dabei spielen könnte, die Motivation im Flow-Zustand während des Spielens von Lernspielen aufrechtzuerhalten. Darüber hinaus könnte sie einen Einfluss auf das deklarative Wissen haben. Auch die einzelnen Komponenten des Flow-Erlebens (Tabelle 2) wurden in Zusammenhang mit der Selbstwirksamkeitserwartung untersucht. In einer Studie mit 102 Teilnehmenden analysierten Peifer et al. (2020) die Wirkung des Feedbacks auf die

Selbstwirksamkeitserwartung beim Lösen von Mathematikaufgaben an einem Computer. Dabei konnten sie zeigen, dass die Selbstwirksamkeitserwartung in Bezug auf das Flow-Erleben als Mediator fungiert (Peifer et al., 2020). In der Untersuchung wurde außerdem ein positiver Effekt auf die Leistung nachgewiesen, der jedoch nicht signifikant war, was auf weitere wirkende Mechanismen schließen lässt (ebd.). Diese Wissenschaftler interpretieren das Ergebnis so, dass Menschen mit einer hohen Selbstwirksamkeitserwartung bessere Leistungen erbringen und deshalb mehr Flow erleben (ebd.). Die Vergleichsstudie von Hung et al., s. Abschnitt 2.2.1) aus Abschnitt 2.2.1. konnte in der Experimentalgruppe ermitteln, dass die naturwissenschaftsbezogenen und technologiebezogenen Selbstwirksamkeitserwartungen stärkere Prädiktoren des Flow-Erlebens als in der Kontrollgruppe waren. Diese Ergebnisse deuten darauf hin, dass Herausforderungen zwar Angst verursachen können, die aber reduziert wird, wenn sich Lernende aktiv und mit Freude an Aufgaben beteiligen.

Die Studie von Rachmatullah et al. (2021) mit einem Serious Game zum Thema *Genetik* umfasste eine Stichprobe von 394 Schüler:innen der Schuljahre 7–12. Dabei wurde unter anderem die gegenseitige Beeinflussung der Variablen Lernwirksamkeit, Flow-Erleben und der Selbstwirksamkeitserwartung in Strukturgleichungsmodellen untersucht. Sie ermittelten hohe Effektstärken ($\eta^2 = .35$) bei der Lernwirksamkeit durch das Serious Game und einen positiven Einfluss einer hohen Selbstwirksamkeitserwartung auf die Lernwirksamkeit. Das Vorwissen korreliert signifikant mit der Lernwirksamkeit ($r = .44$). Ein schwacher, aber signifikanter Zusammenhang konnte auch zwischen Fachwissen und Flow-Erleben ermittelt werden. Auffällig ist, dass die Ergebnisse auch in diesem Fall keinen signifikanten Zusammenhang zwischen Vorwissen und Fachwissen belegen¹². Unabhängig von ihrem Vorwissen schienen die Schüler:innen durch die Maßnahmen des Spiels trotzdem zu wissen, wie sie vorgehen müssen. Darüber hinaus konnte ermittelt werden, dass eine geringe Selbstwirksamkeitserwartung mit Frustrationsgefühlen und einem vermutlich dadurch verminderten Flow-Erleben einhergeht (ebd.).

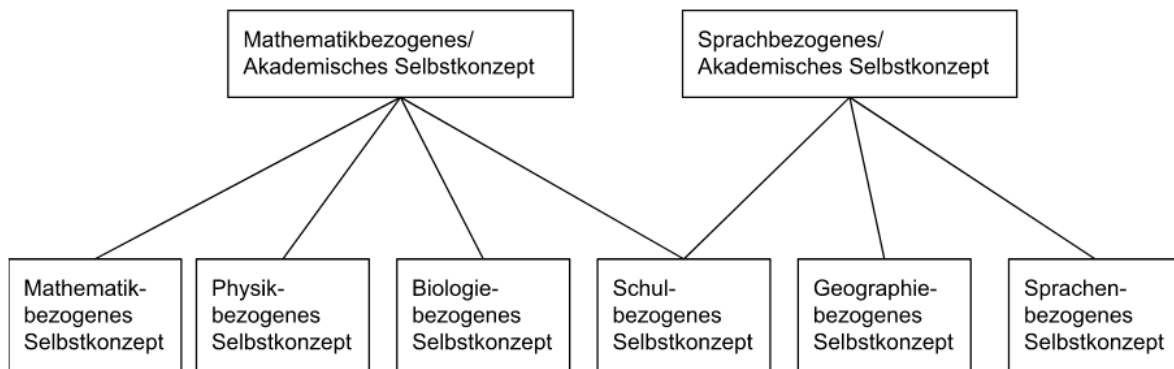
Insgesamt deuten Ergebnisse aus bisherigen Untersuchungen darauf hin, dass Schüler:innen ihre naturwissenschaftsbezogenen Selbstwirksamkeitserwartungen durch Serious Games positiv verändern können. Diese Verbesserung erfolgt trotz geschlechtsspezifischer Unterschiede in Bezug auf die Erfahrung und Einstellung zu Computerspielen (Bergey et al., 2015; Paraskeva et al., 2010). Mit ihren motivationalen Elementen sowie dem Belohnungs- und Feedbacksystem können Serious Games nicht nur auf das Flow-Erleben, sondern auch auf die Selbstwirksamkeitserwartung Einfluss nehmen (X. Yang et al., 2021). Diese wird zusätzlich verbessert, indem in Serious Games Banduras Quellen der Selbstwirksamkeitserwartung Berücksichtigung finden. Die Bewältigung virtueller Aufgaben und die Beherrschung von Inhalten in der virtuellen Umgebung können zu einer Steigerung der Selbstwirksamkeitserwartung in der realen Welt führen (Hardin et al., 2014). Die Selbstwirksamkeitserwartung hat durch ihren Einfluss auf die Lernmotivation eine große Bedeutung im schulischen Kontext. Aber auch andere Faktoren, wie das Selbstkonzept, können die Motivation vorhersagen und beeinflussen.

¹² Vorwissen bezieht sich auf die bereits vorhandenen Kenntnisse und Fertigkeiten einer Person in einem bestimmten Bereich, während Fachwissen spezifisches, detailliertes Wissen in einem bestimmten Fachgebiet ist, das aufgrund seiner kodierten Form auf zahlreichen Medien gespeichert, verarbeitet und übertragen werden kann (König et al. 2018)

2.4 Die Rolle des Selbstkonzepts in Zusammenhang mit Serious Games

Im Allgemeinen¹³ beschreibt das Selbstkonzept das Selbstbild einer Person und deren Annahmen und Überzeugungen in Bezug auf sich selbst (Bracken, 1996). Die grundlegende Definition des Selbstkonzepts für die Forschung bildet die Zusammenfassung von Shavelson et al. (1976) im sog. Shavelson-Modell des Selbstkonzepts (Marsh & Hattie, 1996). Demnach umfasst das Selbstkonzept sieben Merkmale: Es ist strukturiert, mehrdimensional, hierarchisch, stabil, evaluativ, differenziert sich im Laufe des Heranwachsens und ist von anderen Konstrukten abgrenzbar (Shavelson et al., 1976). Diese sieben Merkmale werden im Folgenden kurz erläutert: *Strukturiert* bedeutet, dass Individuen das selbstbezogene Wissen in unterschiedliche Dimensionen strukturieren. Diese *Mehrdimensionalität* stellt nach Shavelson et al. (1976) ein weiteres Merkmal dar. Die Dimensionen sind *hierarchisch* angeordnet, mit dem allgemeinen oder globalen Selbstkonzept an der Spitze, das sich in akademisches, soziales, physisches und emotionales Selbstkonzept untergliedert. Das akademische Selbstkonzept teilt sich weiter in domänenspezifische Selbstkonzepte auf, wie naturwissenschaftliche und technische Selbstkonzepte, und in die sozialen Selbstkonzepte in situationsspezifischen Einschätzungen. Bei empirischen Prüfungen konnte die Mehrdimensionalität bestätigt werden (Marsh, 1990; Shavelson & Bolus, 1982). Insbesondere das akademische Selbstkonzept wurde als eigene Dimension identifiziert (Hattie, 1992; Marsh, 1990). Dabei war eine aufschlussreiche Erkenntnis, dass zwischen dem Mathematik- und dem Sprach-Selbstkonzept nur eine geringe Korrelation besteht, obgleich beide Unterkategorien des akademischen Selbstkonzepts sind und sich daher höhere Korrelationen hätten zeigen sollen (Marsh & Craven, 2006; Marsh et al., 2019). Dies führte zu einer ersten Überarbeitung des Marsh-Shavelson-Modells (siehe Abbildung 1), bei dem das globale Selbstkonzept direkt in ein mathematikbezogenes-akademisches und ein sprachbezogenes-akademisches Selbstkonzept aufgeteilt wurden (Marsh et al., 1988).

Abbildung 1 Marsh-Shavelson-Modell des Selbstkonzepts (übersetzt nach Marsh et al., 1988).



Neuere Selbstkonzeptstudien mit einem Fokus auf domänenspezifischen Selbstkonzepten haben gezeigt, dass auch Globalität nicht unbedingt in der Konstruktdefinition enthalten ist (Byrne, 1996). In der Selbstkonzeptforschung wird vermutet, dass die Entwicklung des Selbstkonzepts zwar einen lebenslangen Prozess darstellt (Mummendey, 2006), jedoch das Jugendalter eine entscheidende Phase für die *Herausbildung* insbesondere der bereichsspezifischen akademischen Selbstkonzepte ist. Während das globale Selbstkonzept *stabil* ist und eine Änderung mehrerer Interventionen bedarf, sind

¹³ Die Definition des Konstrukts „Selbstkonzept“ hat eine lange Geschichte, unterschiedliche Theoretiker haben ihre Ideen zusammengetragen und es gibt nicht den Urheber, was zu entsprechend vielen Operationalisierungen und sogar Namen führte (Pajares & Schunk, 2006).

die einzelnen Dimensionen einfacher zu beeinflussen. Es wird vermutet, dass die Stabilität mit absteigender Hierarchie der Dimensionen abnimmt (Shavelson et al., 1976). Jedoch konnte diese Vermutung bisher nicht abschließend empirisch bestätigt werden (Marsh & Hattie, 1996; Shavelson & Bolus, 1982). *Evaluative* Erfahrungen mit der Umwelt beeinflussen und prägen das Selbstkonzept, wobei der soziale Vergleich, insbesondere mit wichtigen anderen Personen, als besonders bedeutender Faktor bei der Entwicklung gilt (Shavelson et al., 1976). Sein entscheidender Einfluss auf die Bildung des akademischen Selbstkonzepts konnte empirisch nachgewiesen werden (Coleman & Fults, 1982). Die Eigenbewertung wird teilweise als affektiv-evaluativer Teil des Selbstkonzepts beschrieben, sodass sich frühere Erfolge und Misserfolge auf spätere Selbstkonzepte auswirken (Möller & Trautwein, 2009; Skaalvik, 1997). Die Bewertung der eigenen Kompetenzen (Möller & Köller, 2004) ist ein weiterer Parameter, bei dem vermutet wird, dass er einen besonders signifikanten Einfluss auf die Motivation und das Lernen von Schülern hat (Harter, 1988).

Es wird weiterhin angenommen, dass die Motivation ein Mediator zwischen Selbstkonzept und schulischer Leistung ist (Hellmich & Günther, 2011). Die *Abgrenzung von anderen Konstrukten*, wie der Selbstwirksamkeitserwartung, ist die letzte Annahme des Selbstkonzepts. Selbstkonzept und Selbstwirksamkeitserwartung erhielten wegen ihrer Vorhersage von Motivation und schulischer Leistungen der Schüler:innen viel Beachtung und waren lange Zeit mit dem Mangel an klaren differenzierten wörtlichen Bedeutungen und entsprechend differenzierten operativen Definitionen von Selbstkonstrukten konfrontiert (Wylie, 1968). Vermutlich sind das akademische Selbstkonzept und die Selbstwirksamkeitsüberzeugungen – durch ihre hierarchische Struktur – auf einer spezifischen Ebene des selben Bereichs nicht voneinander zu trennen (Pajares, 1997), weshalb ein besonderes Augenmerk auf die Wahl geeigneter Skalen zur Erfassung der Konstrukte liegen sollte. Bei Items, die die Selbstwirksamkeitserwartung erfassen, sollten Proband:innen nicht dazu aufgefordert werden, ihre Fähigkeiten mit denjenigen anderer Personen zu vergleichen. Stattdessen sollen sie ihre Kompetenz zielgerichtet anhand eines Bezugspunktes beurteilen (Bong & Skaalvik, 2003). Items zum Selbstkonzept beruhen vor allem auf Informationen aus sozialen Vergleichen, wobei die eigenen Fähigkeiten eingeordnet werden (ebd.). Insgesamt ist ein größerer Vorhersagenutzen durch die akademische Selbstwirksamkeitserwartung zu vermuten, da der Schwerpunkt auf den eigenen früheren Erfahrungen in Zusammenhang mit der Bewältigung von Aufgaben liegt. Mit dem Selbstkonzept können Angst und Besorgnis, aber auch Motivation, vor allem im Kontext des Lernens, vorhergesagt werden (Bandalos et al., 1995; Hellmich & Günther, 2011; Pajares & Miller, 1994; Skaalvik & Rankin, 1995). Insbesondere Letzteres ist für diese Arbeit von Interesse.

Auch in Zusammenhang mit Serious Games wird vermutet, dass die spezifischen Selbstkonzepte Prädiktoren der Motivation sind. Es wird angenommen, dass Schüler:innen mit einem hohen computerbezogenen Selbstkonzept stärker motiviert sind, sich mit dem Spiel auseinanderzusetzen. Demgegenüber beschäftigen sich Schüler:innen mit einem hohen physikbezogenen Selbstkonzept bevorzugt mit den Inhalten, die ihren Interessen oder Kompetenzen entsprechen. Hierbei wird von geschlechtsspezifischen Unterschieden des Selbstkonzepts ausgegangen, sowohl in Bezug auf die Physik (D. H. Rost & Sparfeldt, 2002) als auch auf den Umgang mit Computern (Janneck et al., 2012). In früheren Studien gaben Frauen und Mädchen an, weniger Erfahrung im Umgang mit Computern zu haben (Broos & Roe, 2006; Eickelmann et al., 2019; L. J. Nelson & Cooper, 1997; Whitley, 1997). Zudem empfanden sie im Vergleich zu Jungen und Männern diese Erfahrungen seltener als positiv (Campbell, 1990; Whitley, 1997). Während Frauen die Computer oft aus pragmatischen Gründen benutzen und einen erfolgreichen Gebrauch auf externe Ursachen wie Glück zurückführen, gaben Männer als Grund für die Nutzung die intrinsische Motivation an (Campbell, 1990; L. J. Nelson & Cooper, 1997; Wetzel, 2002). Obgleich sich bei den Jugendlichen mittlerweile die Differenzen zwischen den Geschlechtern

hinsichtlich der Kompetenz bei der Computernutzung verringert haben (Eickelmann et al., 2019), schätzen Jungen und Männer ihre Fähigkeiten nach wie vor höher ein und sind aufgeschlossener im Umgang mit neuen Computeranwendungen.

Ein hohes computerbezogenes Selbstkonzept kann demnach mit einer starken Motivation bezüglich des Serious Games einhergehen und durch die intensivere Interaktion mit dem Medium die Lernergebnisse stärken (Eickelmann et al., 2019). Darüber hinaus gibt es Erkenntnisse, dass Frauen, die Videospiele mögen, tendenziell mehr an Mathematik und Naturwissenschaften interessiert sind als Nichtspielerinnen (Giammarco et al., 2015). Jedoch konnten keine Zusammenhänge zwischen der Notenleistung und dem Selbstkonzept ermittelt werden (Häußler & Hoffmann, 1995; Schiepe-Tiska et al., 2016), sodass der Einfluss auf das Lernergebnis vermutlich gering ist. Bisher haben sich wenige Forschungsarbeiten mit der Veränderung des Selbstkonzepts durch Serious Games auseinandergesetzt, jedoch erscheint eine kurzfristige Maßnahme wie das Spielen eines einzelnen Serious Games auch nicht sinnvoll. Aufgrund der vermuteten Stabilität des Selbstkonzepts ist davon auszugehen, dass nur durch mehrere und langfristige Maßnahmen Veränderungen erzielt werden können. Erste Bestätigungen für diese Annahme finden sich bspw. bei Ertl (2016). Der Einsatz eines Serious Games zum Thema *Entrepreneurship* mit physikbezogenen Elementen bewirkte in einem Pre-Post-Design mit 35 Schüler:innen zwischen 13 und 14 Jahren eine Steigerung des Selbstkonzepts, jedoch war diese nicht signifikant (ebd.). Auch Kapp et al. (2019) konnten in ihrer Untersuchung mit einem technikbezogenen Serious Game eine Änderung des Selbstkonzepts feststellen ($M_{pre} = 4.09$, $SD_{pre} = .79$; $M_{post} = 4.10$, $SD_{post} = .80$), die jedoch auch hier nicht signifikant ausfiel (ebd.). Die Zusammenhänge zwischen Geschlecht, Selbstkonzept, Flow-Erleben, Selbstwirksamkeitserwartung und Lernwirksamkeit werden im nächsten Abschnitt in einem nomologischen Netzwerk zusammengetragen.

2.5 Einordnung der Konstrukte in ein nomologisches Netzwerk

Ein nomologisches Netzwerk ist ein theoretischer Rahmen, der die Beziehungen zwischen einer Reihe von Konzepten oder Variablen beschreibt. Es soll erklären, wie die folgenden Variablen miteinander in Beziehung stehen: Geschlecht, computerbezogenes und physikbezogenes Selbstkonzept, Flow-Erleben, Fachwissen in der Elektrizitätslehre und physikbezogene Selbstwirksamkeitserwartung. In Abbildung 2 sind die vermuteten Pfade dargestellt, die im Folgenden hergeleitet werden. Die Buchstaben in Klammern bezeichnen den jeweiligen Pfad und finden sich auch in der Abbildung 2 wieder.

Die theoretischen Herleitungen der letzten Abschnitte lassen die Vermutung zu, dass das Geschlecht auf alle genannten Variablen einen Einfluss hat (a). So konnten bereits geschlechtsspezifische Unterschiede im computerbezogenen (Janneck et al., 2012) und physikbezogenen (D. H. Rost & Sparfeldt, 2002) Selbstkonzept gezeigt werden. Sie haben möglicherweise auch Auswirkungen auf die intrinsische Motivation, sich mit physikalischen Inhalten zu beschäftigen oder sich auf ein Computerspiel einzulassen (Whitton, 2007). Aufgrund der Forschungsergebnisse zu Geschlechterunterschieden bei der Selbstwirksamkeitserwartung wird auch hier von einem Einfluss ausgegangen (Kosuch, 2010). Bedingt durch diese Unterschiede ist auch von einem geringeren Interesse am Physikunterricht seitens der Schülerinnen auszugehen, was zu einem Unterschied im Vorwissen in der Elektrizitätslehre zum Zeitpunkt t_1 vor der Erhebung (im Folgenden Fachwissen(t_1)) führen könnte (Bartosch, 2013). Andererseits zeigen Studien, dass Mädchen in Physik identische Leistungen erbringen wie Jungen (OECD, 2016), sodass auch möglich ist, dass hier kein Unterschied ist.

Eine positive Wahrnehmung der eigenen Fähigkeiten könnte eher in einen Flow-Zustand versetzen, während eine negative Wahrnehmung das Gegenteil bewirken könnte, sodass auch das Flow-Erleben nicht nur durch das Fachwissen(t_1) sondern darüber auch durch das Geschlecht beeinflusst sein könnte (Engeser & Vollmeyer, 2005).

Da es sich bereits gezeigt hat, dass das Selbstkonzept die Motivation vorhersagen kann, wird erwartet, dass über das computerbezogene (b) und das physikbezogene (c) Selbstkonzept eine Prognose des Flow-Erlebens gegeben werden kann. Der Grund für diese Annahme ist zum einen, dass sich Kinder mit einem positiven computerbezogenen Selbstkonzept gerne mit dem Serious Game auseinandersetzen, und zum anderen, dass sich Kinder mit einem positiven physikbezogenen Selbstkonzept gerne mit Physik-Inhalten befassen, was der ersten Komponente . Es wird angenommen, dass diese positiven Selbstkonzepte sowohl die Motivation als auch das Potenzial für Flow-Erlebnisse erhöhen können.

Theoriegeleitet kann darauf geschlossen werden, dass auch das Fachwissen(t_1) sich auf das Flow-Erleben auswirkt (d). da eine Passung von Fachwissen(t_1) und geforderten Kompetenzen als eine der Komponenten bzw. Voraussetzungen in Tabelle 2 genannt wird (Engeser & Rheinberg, 2008; Engeser et al., 2005). Jedoch konnte in Studien bisher nicht immer ein signifikanter Zusammenhang zwischen dem Fachwissen(t_1) und dem Flow-Erleben ermittelt werden (Rachmatullah et al., 2021).

Die Bedingungen, um Flow zu erleben, wie z. B. das Erhalten von Feedback, stehen auch in einer möglichen Verbindung mit den Quellen der Selbstwirksamkeitserwartung (Peifer, 2017). Beide Konstrukte stehen in einer Verbindung mit Motivation und Leistung und können dazu beitragen, dass Menschen sich engagierter, kompetenter und zufriedener bei einer Tätigkeit fühlen (ebd.).

Der Einfluss des Selbstkonzepts auf die Lernleistung konnte bereits gezeigt werden (Köller et al., 2006). Auch die großen Schulstudien wie PISA (OECD, 2016) und TIMMS (Kleickmann et al., 2012) stellten eine höhere Leistungsbereitschaft bei Schüler:innen mit höherem Selbstkonzept in den Naturwissenschaften fest. Allerdings werden diese Einschätzungen auch durch Leistungen in anderen Fächern, Genderstereotype und Vergleiche mit Peers beeinflusst (Wolff et al., 2018). Für diese Arbeit wird angenommen, dass sich Personen mit starken Selbstkonzepten (f, g) intensiver mit den Inhalten des Serious Games auseinandersetzen und in der Folge ein höheres Ergebnis beim Fachwissen nach dem Spiel (im folgenden Fachwissen(t_2)) erzielen (Hellmich & Günther, 2011).

Faktoren wie Motivation und Flow-Erleben, gelten aus denselben Gründen wie die Selbstkonzepte und die intensivere Auseinandersetzung mit den Inhalten des Serious Games als zentrale Voraussetzung für das Lernen (Pekrun, 2011). Daher wird davon ausgegangen, dass sich der erlebte Flow (j) im Serious Game positiv auf das Fachwissen(t_2) auswirkt (Csikszentmihalyi & LeFevre, 1989; Csikszentmihalyi & Schiefele, 1993; Kiili, 2005; Moneta & Csikszentmihalyi, 1999; Skadberg & Kimmel, 2004).

Der direkte Zusammenhang zwischen Fachwissen(t_1) und Fachwissen(t_2) (l) könnte sich als nicht linear herausstellen, was darin begründet liegt, dass Schüler:innen mit viel Vorwissen durch das Spiel keinen großen Lernzuwachs erleben (Virvou et al., 2005). Möglicherweise wird die Stärke der Beziehung durch eine Moderatoreigenschaft des Flow-Erlebens beeinflusst und Schüler:innen mit hohem Vorwissen haben einen höheren Lernzuwachs.

Ob ein direkter Einfluss des Fachwissens(t_1) auf die Selbstwirksamkeitserwartung(t_2) (m) besteht, ist fraglich. Einerseits ist ein Zusammenhang zwischen erlebtem Erfolg im Serious Game, der durch Vorwissen bedingt ist, wahrscheinlich und könnte sich positiv auf die Selbstwirksamkeitserwartung auswirken. Andererseits konnte in Studien wie PISA (OECD, 2016) festgestellt werden, dass auch Schüler:innen mit guten schulischen Leistungen niedrige Selbstwirksamkeitserwartungen aufweisen. Kein Zusammenhang wird zwischen der Selbstwirksamkeitserwartung(t_1) und dem Fachwissen(t_2) (o) vermutet, da ermittelt wurde, dass sich Schüler:innen mit geringer und mit hoher

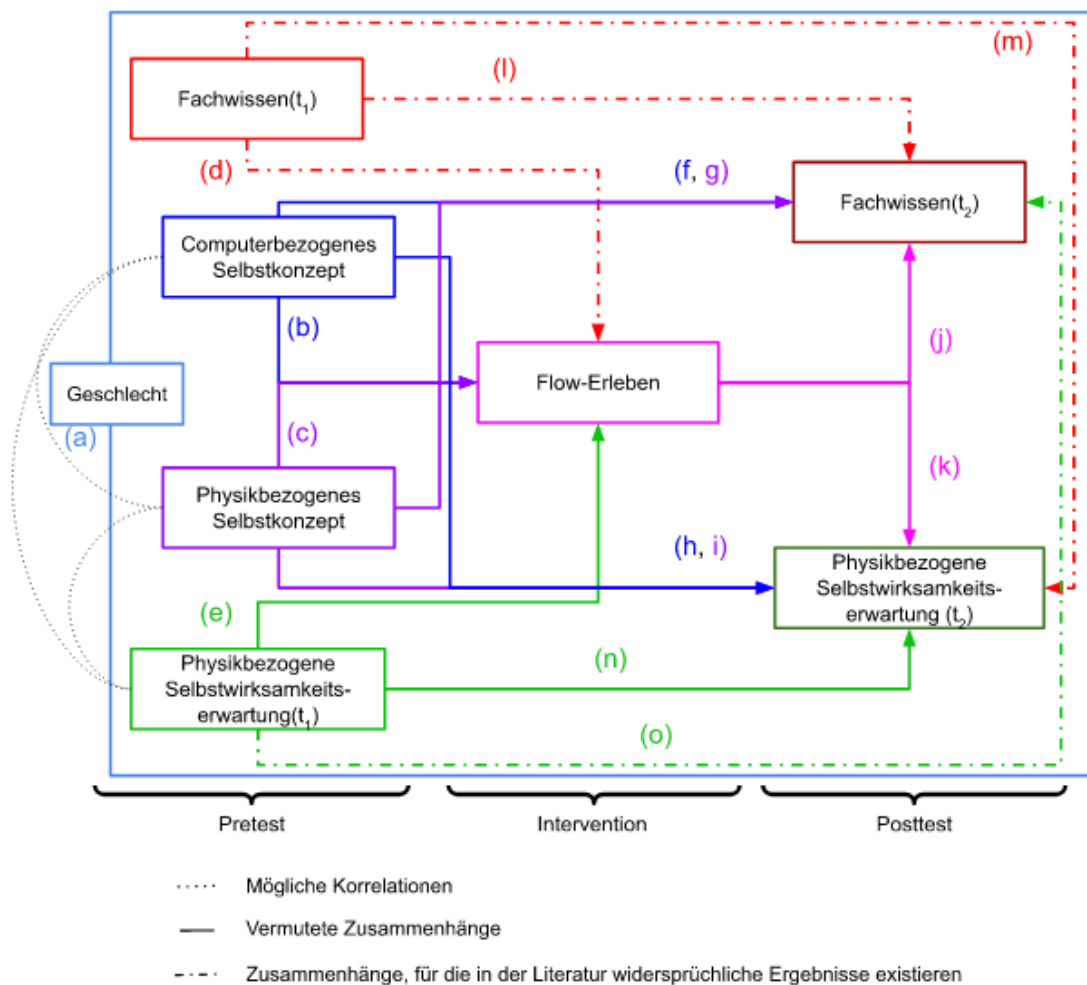
Selbstwirksamkeitserwartung gleich intensiv mit Lerninhalten auseinandersetzen (Ketelhut, 2007). Jedoch wird von einem starken Zusammenhang zwischen der Selbstwirksamkeitserwartung(t_1) und der Selbstwirksamkeitserwartung(t_2) ausgegangen (n).

Eine Beziehung zwischen den Selbstkonzepten und der Selbstwirksamkeitserwartung(t_2) kann aufgrund der möglichen Ähnlichkeit der Konstrukte ebenfalls angenommen werden (Marsh et al., 2019). So kommen Pajares und Schunk (2006) zu dem Ergebnis, dass Schüler:innen mit einem positiven Selbstkonzept neue Aufgaben mit hoher Selbstwirksamkeitserwartung angehen. Bestätigt wurde dies durch eine Pfadanalyse auf Grundlage der belgischen PISA-Daten des Jahres 2003 durch Ferla et al. (2009). Somit kann das akademische Selbstkonzept als guter Prädiktor für die Selbstwirksamkeitserwartung(t_2) angenommen werden (h,i).

Zwischen Selbstwirksamkeitserwartung und dem erlebten Flow existiert möglicherweise eine Wechselwirkung, sodass eine hohe Selbstwirksamkeitserwartung(t_1) vermutlich das Flow-Erleben vorhersagen kann (e) und diese wiederum die Selbstwirksamkeitserwartung(t_2) Selbstwirksamkeitserwartung (Peifer, 2017). Die Beziehung zwischen Selbstwirksamkeitserwartung und Flow-Erleben kann möglicherweise durch eine Mediatorfunktion erklärt werden., bei der das Flow-Erleben die Selbstwirksamkeitserwartung zu zwei Zeitpunkten vermittelt.

Die Zusammenhänge des nomologischen Netzwerkes sollen in dieser Arbeit empirisch anhand der Fragestellungen des nächsten Kapitels überprüft werden.

Abbildung 2 Nomologisches Netzwerk der theoretischen Zusammenhänge.



3 Fragestellungen und Hypothesen

Unter Berücksichtigung der Zusammenhänge, die theoretisch hergeleitet und im nomologischen Netzwerk dargelegt wurden, sollen in der vorliegenden Arbeit die Wirkungen eines Serious Games auf das Fachwissen und die Selbstwirksamkeitserwartung empirisch untersucht werden. Es wird angenommen, dass vor allem die intrinsische Motivation in Serious Games dabei eine zentrale Funktion einnimmt. Die intrinsische Motivation wird dabei im Zusammenhang mit der Flow-Theorie angenommen. In einer Intervention sollen Schüler:innen der 8. Klasse zufällig in eine Experimentalgruppe und eine Kontrollgruppe eingeteilt werden. Durch die Durchführung einer Intervention mit einer Experimentalgruppe und einer Kontrollgruppe kann die Wirksamkeit von Serious Games im Vergleich zu konventionellen Unterrichtsmethoden direkt gemessen und mögliche Auswirkungen von Serious Games auf das Fachwissen und die Selbstwirksamkeitserwartung empirisch untersucht werden. Die Experimentalgruppe spielt ein Serious Game zum Thema *Elektrizitätslehre* und die Kontrollgruppe erhält konventionelle Unterrichtsmaterialien wie Arbeitsblätter und Versuchsmaterial. Eine detaillierte Beschreibung des Aufbaus der Intervention folgt in Kapitel 4.

Zunächst sollen in der Stichprobe geschlechtsspezifische Unterschiede untersucht werden. Bedingt durch die jeweilige Erfahrung im Umgang mit Computern (Beierlein, 2011) ist von einem Unterschied auszugehen, der sich bereits in den Prädiktorvariablen zeigen wird. Ebenfalls in Zusammenhang mit physikbezogenen Einstellungen konnten Unterschiede zwischen den Geschlechtern ermittelt werden (OECD, 2016). Daher wird davon ausgegangen, dass Schülerinnen in den physikbezogenen Selbstkonzepten und der physikbezogenen Selbstwirksamkeitserwartung unter den Werten der Schüler liegen. In der ersten Forschungsfrage geht es um die Unterschiede zwischen den beiden Interventionsgruppen.

1. Profitieren Schüler:innen, die ein Serious Game zum Thema *Elektrizitätslehre* spielen, stärker hinsichtlich des Flow-Erlebens, des Fachwissens zur Elektrizitätslehre und der physikbezogenen Selbstwirksamkeitserwartung als Schüler:innen in einem konventionellen Unterrichtsetting?

Aufgrund bisheriger Forschungsergebnisse wird ein höheres Flow-Erleben in der Experimentalgruppe erwartet, unter anderem weil das Serious Game die aus der Wissenschaft bekannten Flow-förderlichen Elemente berücksichtigt (siehe Unterabschnitte 2.2.2 und 4.1.2). Ausgehend von der Hypothese, dass das Serious Game den Flow-Zustand ermöglicht, ist zu untersuchen, ob in der Experimentalgruppe die Lernwirksamkeit höher ist als in der Kontrollgruppe. Da alle Quellen der Selbstwirksamkeitserwartung im Serious Game berücksichtigt werden, ist anzunehmen, dass die physikbezogene Selbstwirksamkeitserwartung in der Experimentalgruppe stärker gefördert wird als in der Kontrollgruppe.

Zwar besteht die Möglichkeit, dass Schüler aufgrund ihrer Neigung zum Fach Physik und zu Computerspielen höhere Ergebnisse beim Flow-Erleben erzielen als Schülerinnen. Da aber die Komponenten des Flow-Erlebens im Spiel Berücksichtigung finden, wird dieser Unterschied eventuell dadurch ausgeglichen. Daher werden beim erlebten Flow keine Differenzen zwischen den Geschlechtern in der Experimentalgruppe erwartet, jedoch in der Kontrollgruppe. Aufgrund der bisherigen theoretischen Befunde, unter anderem auch aus den Leistungsstudien (OECD, 2016), wird davon ausgegangen, dass es bei den Lernergebnissen keinen Unterschied zwischen den Geschlechtern gibt. Die Annahme beruht auch darauf, dass das Serious Game so gestaltet wurde, dass es insbesondere Mädchen anspricht. Ergänzend soll daher die Hypothese geprüft werden, ob in der Kontrollgruppe Unterschiede zwischen den Geschlechtern existieren. Dabei wird in Anlehnung an bisherige Befunde (ebd.) davon ausgegangen, dass es auch hier keine Unterschiede zwischen den

Geschlechtern gibt. Aufgrund der vermuteten Unterschiede in der Selbstwirksamkeitserwartung, bei der Schülerinnen gegenüber Schülern geringere Werte in der Selbstwirksamkeitserwartung aufweisen und diese im Laufe der Zeit tendenziell abnimmt (Solga & Pfahl, 2009), soll untersucht werden, ob bezüglich der Selbstwirksamkeitserwartung die Schülerinnen im gleichen Maße vom Serious Game profitieren können wie die Schüler. Für die Kontrollgruppe wird vermutet, dass vor allem Schüler bezüglich der Selbstwirksamkeitserwartung profitieren. Die Überlegungen finden sich in den Hypothesen in Tabelle 4 wieder.

Die zweite Forschungsfrage betrifft die Voraussetzungen, um Flow zu erleben.

2. Inwiefern beeinflussen Schüler:innen-bezogene Merkmale wie Geschlecht, computerbezogenes und physikbezogenes Selbstkonzept, Fachwissen in der Elektrizitätslehre und physikbezogene Selbstwirksamkeitserwartung das Flow-Erleben?

In der Experimentalgruppe wird kein Unterschied bezüglich des erlebten Flows zwischen den Geschlechtern vermutet. Daher ist davon auszugehen, dass die Prädiktoren unterschiedlich wirken und sich möglicherweise auch, basierend auf dem Geschlecht, eine Vorhersage treffen lässt, ob Flow erlebt wird. Auch das computerbezogene und das physikbezogene Selbstkonzept können, wie bereits hergeleitet, Prädiktoren des Flow-Erlebens sein. Das physikbezogene Selbstkonzept, bei dem es um die Bewältigungserfahrungen aus dem Physikunterricht geht, ist möglicherweise sowohl in der Experimentalgruppe als auch in der Kontrollgruppe ein Prädiktor für das Flow-Erleben. Das computerbezogene Selbstkonzept wird als Prädiktor für das Flow-Erleben durch Serious Games angenommen. Dieser wird vermutlich nur für die Experimentalgruppe einen Zusammenhang zeigen, nicht jedoch in der Kontrollgruppe, da in dieser keine Interaktion mit Computern stattfindet, die motivierend wirken kann. Ein weiteres Konstrukt, das für die Erklärung der Motivation in Serious Games von Bedeutung sein könnte, ist die Selbstwirksamkeitserwartung (siehe Abschnitt 2.3). Es soll geprüft werden, ob die Selbstwirksamkeitserwartung in der Experimentalgruppe ein Prädiktor ist. In Studien konnte nachgewiesen werden, dass Serious Games auch bei unterschiedlicher Ausprägung der Selbstwirksamkeitserwartung durch ihre Elemente einen ausgleichenden Effekt bezüglich des Flow-Erlebens zeigen. Da vermutet wird, dass das Erleben von Flow in der Kontrollgruppe eine hohe Motivation voraussetzt, wird auch hier erwartet, dass ein Zusammenhang zwischen Flow-Erleben und Selbstwirksamkeitserwartung besteht. Das Vorwissen ist bei der Experimentalgruppe keine Voraussetzung, um Flow zu erleben, da das Spiel durch seine Hinweise und Hilfestellungen die Schüler:innen unterstützt. Vermutlich zeigt sich jedoch ein Zusammenhang zwischen dem Vorwissen und dem Flow-Erleben für die Kontrollgruppe, da vermutet wird, dass Schüler:innen mit Vorwissen auch stärker an Aufgaben im Kontext Physik interessiert sind, sodass hier die Komponente 1 (Tabelle 2) Passung der Kompetenzen mit der Herausforderung Flow-Erleben zu ermöglichen eine wichtigere Rolle spielt (siehe Abschnitt 2.2). Die Überlegungen zur zweiten Fragestellung werden als Hypothesen in der Tabelle 5 aufgegriffen.

Mit der dritten Fragestellung werden die Zusammenhänge aller Variablen in einem Serious Game mit dem Fachwissen in der Elektrizitätslehre untersucht.

3. Welche Wirkungen hat das Flow-Erleben auf das Fachwissen zur Elektrizitätslehre und die physikbezogene Selbstwirksamkeitserwartung?

Wie in Abschnitt 2.2 dargestellt, wird Flow-Erleben mit einer erhöhten Persistenz beim Lernen in Verbindung gebracht, was möglicherweise die Lernergebnisse und die Selbstwirksamkeitserwartung beeinflusst (Ke, 2009; Seelhammer & Niegemann, 2009). In einem Strukturgleichungsmodell sollen die Pfade des gesamten nomologischen Netzwerks überprüft werden. Dabei wird vermutet, dass das Flow-Erleben eine Moderatorrolle bei möglichen Beziehungen zwischen Fachwissen(t_1) und Fachwissen(t_2) und eine Mediatorrolle zwischen Selbstwirksamkeitserwartung(t_1) und

Selbstwirksamkeitserwartung(t_2) spielt. Die Hypothesen zu dieser Fragestellung finden sich in der Tabelle 6. Die folgenden Tabellen (Tabelle 4–Tabelle 6) geben eine Übersicht über die den Forschungsfragen zugeordneten Hypothesen, die in dieser Arbeit untersucht werden.

Tabelle 4 *Hypothesen zur ersten Fragestellung.*

Nr.	Alternativhypothese
H1.1	In der Experimentalgruppe mit einem Serious Game wird mehr Flow erlebt als in der Kontrollgruppe.
H1.2	Die Schüler:innen der Experimentalgruppe profitieren stärker in Bezug auf das Fachwissen als die Schüler:innen der Kontrollgruppe.
H1.3	Die Schüler:innen der Experimentalgruppe profitieren in Bezug auf die Selbstwirksamkeitserwartung stärker als die Schüler:innen der Kontrollgruppe.
H1.4	Innerhalb der Experimentalgruppe gibt es bezüglich des erlebten Flows keine Unterschiede zwischen den Geschlechtern.
H1.5	Innerhalb der Kontrollgruppe erleben die Schüler mehr Flow als die Schülerinnen.
H1.6	Innerhalb der Experimentalgruppe gibt es bezüglich der Verbesserung des Fachwissens keine Unterschiede zwischen den Geschlechtern.
H1.7	Innerhalb der Kontrollgruppe gibt es bezüglich der Verbesserung des Fachwissens keine Unterschiede zwischen den Geschlechtern.
H1.8	Innerhalb der Experimentalgruppe gibt es bezüglich der Selbstwirksamkeitserwartung keine Unterschiede zwischen den Geschlechtern.
H1.9	Innerhalb der Kontrollgruppe profitieren die Schüler bezüglich der Selbstwirksamkeitserwartung stärker als die Schülerinnen.

Tabelle 5 Hypothesen zur zweiten Fragestellung.

Nr.	Alternativhypothese
H2.1	Die Interventionsart (Experimentalgruppe oder Kontrollgruppe) hat einen Einfluss auf das Flow-Erleben.
H2.2	Das Vorwissen in der Elektrizitätslehre ist ein Prädiktor für das Flow-Erleben.
H2.3	Das Geschlecht ist in der Experimentalgruppe kein Prädiktor für das Flow-Erleben.
H2.4	Das physikbezogene Selbstkonzept ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.
H2.5	Das computerbezogene Selbstkonzept ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.
H2.6	Die physikbezogene Selbstwirksamkeitserwartung ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.
H2.7	Das Vorwissen in der Elektrizitätslehre ist in der Experimentalgruppe kein Prädiktor für das Flow-Erleben.
H2.8	Das Geschlecht ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.
H2.9	Das physikbezogene Selbstkonzept ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.
H2.10	Das computerbezogene Selbstkonzept ist in der Kontrollgruppe kein Prädiktor für das Flow-Erleben.
H2.11	Die physikbezogene Selbstwirksamkeitserwartung ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.
H2.12	Das Vorwissen in der Elektrizitätslehre ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.

Tabelle 6 Hypothesen zur dritten Fragestellung.

Nr.	Alternativhypothese
H3.1	Das Flow-Erleben ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.
H3.2	Das Flow-Erleben ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.
H3.3	Das Flow-Erleben ist ein Moderator für die Veränderung des Fachwissens in der Experimentalgruppe.
H3.4	Das Flow-Erleben ist ein Mediator für die Veränderung der Selbstwirksamkeitserwartung in der Experimentalgruppe.
H3.5	Das physikbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.
H3.6	Das computerbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.
H3.7	Das physikbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.
H3.8	Das computerbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.
H3.9	Das Flow-Erleben ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.
H3.10	Das Flow-Erleben ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.
H3.11	Das Flow-Erleben ist ein Moderator für die Veränderung des Fachwissens in der Kontrollgruppe.
H3.12	Das Flow-Erleben ist ein Mediator für die Veränderung der Selbstwirksamkeitserwartung in der Kontrollgruppe.
H3.13	Das physikbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.
H3.14	Das computerbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.
H3.15	Das physikbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.
H3.16	Das computerbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.

4 Methodenteil

In diesem Kapitel werden zunächst die fachwissenschaftlichen und bildungswissenschaftlichen Grundlagen vorgestellt, auf denen die Intervention im Fach Physik beruht. Dabei wird mit der Präsentation der Interventionsmaterialien der Experimentalgruppe und der Kontrollgruppe begonnen. Zudem werden die in dieser Arbeit eingesetzten Instrumente und ihre Gütekriterien sowie die statistischen Methoden beschrieben, die zur Überprüfung der Hypothesen aus Kapitel 3 benötigt werden. Anschließend werden das Studiendesign und die Ergebnisse der ersten Pilotierung dargelegt.

4.1 Instrumente

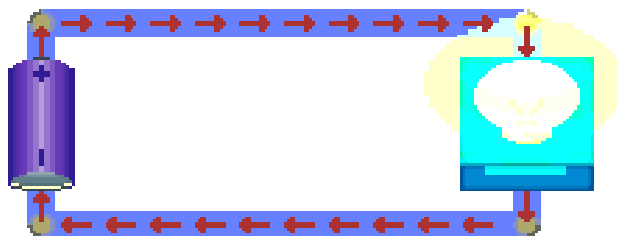
In diesem Abschnitt sollen die Zusammenstellung der Instrumente sowie die eingesetzten Fragebögen erläutert werden. Nach einer kurzen Beschreibung der zugrunde liegenden fachwissenschaftlichen Inhalte der Intervention werden die Interventionsmaterialien der Experimental- und der Kontrollgruppe vorgestellt. Im Anschluss daran erfolgt die Erläuterung der eingesetzten Skalen und der Zusammenstellung der Fragebögen.

4.1.1 Fachwissenschaftliche Inhalte der Intervention

In der Elektrizitätslehre geht es um die Eigenschaften und das Verhalten der physikalischen Größe *Ladung* bzw. *Elektrizität* Q . Dabei wird zwischen den Bereichen *Elektrostatik* und *Elektrodynamik* unterschieden. Während bei der Elektrostatik die Eigenschaften einzelner Ladungen betrachtet werden, liegt der Fokus der Elektrodynamik auf strömenden Ladungen. Diese können im anfänglichen Physikunterricht *Elektrizität* genannt werden. In einem *Leiter* aus Metall kann sich die Ladung frei bewegen und wird durch die *elektrische Spannung* U , der Differenz zwischen der potenziellen elektrischen Energie an den zwei Leiterenden, bewegt. Vereinfacht kann dieser Antrieb der Ladungen im Unterricht als *Elektrizitätsantrieb* bezeichnet werden. Durch die Querschnittsfläche A eines Leiters im geschlossenen Stromkreis tritt in der Zeitspanne Δt eine Ladungsmenge ΔQ . Dieser zeitlich konstante Ladungsfluss ist die *Stromstärke* I .

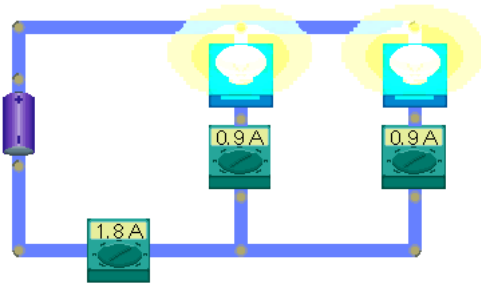
Ein Schaltkreis besteht im einfachsten Fall aus einem Elektrizitätsantrieb, z. B. einer Batterie, einem Akku oder Netzgerät, sowie einem *Nutzer* der Elektrizität, z. B. einer Glühlampe oder einem Motor, die durch elektrische Leiter verbunden werden (siehe Abbildung 3). In einem einfachen elektrischen Stromkreis befindet sich ein Nutzer der Elektrizität sowie ein Elektrizitätsantrieb, und die Elektrizität fließt an jeder Stelle der Leiter gleich stark.

Abbildung 3 Darstellung eines einfachen Schaltkreises.



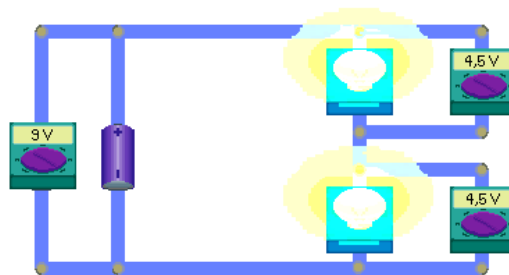
Aus dem Erhaltungssatz für Ladungen ergeben sich die Grundregeln zur Analyse elektrischer Schaltungen, die *Kirchhoffschen Regeln*. Diese besagen, dass die Summe der Ströme an allen Verzweigungspunkten der Leiter, den *Knoten*, einer Schaltung null ergibt: $I_{ges} = 0$ (*Knotenregel*, Abbildung 4).

Abbildung 4 Kirchhoffsche Knotenregel



Zudem ist in einer *Masche*, einem über die Leiter geschlossener Umlauf, die Stromstärke an jedem in Reihe geschalteten Widerstand gleich und so gilt für die Summe der Spannungen an einer Masche: $U_{ges} = 0$ (*Maschenregel*, Abbildung 5).

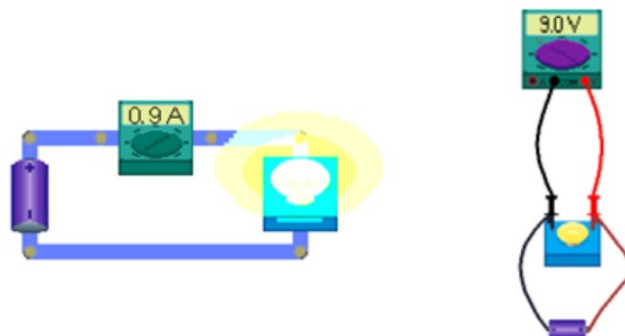
Abbildung 5 Kirchhoffsche Maschenregel.



Für die Intervention wurde die technische Stromrichtung vom Pluspol zum Minuspol angenommen. Sind die beiden Pole eines Elektrizitätsantriebs direkt miteinander verbunden, fließt die Ladung direkt zurück in den Elektrizitätsantrieb. Ohne einen Nutzer, der die Ladung hemmt, wird die Stromstärke sehr groß, sodass es zu einem *Kurzschluss* kommt.

Zur Messung elektrischer Größen werden *Multimeter* eingesetzt. Elektrische Stromstärken werden so erhoben, dass das Multimeter in den zu messenden Leitungsabschnitt eingebaut wird, sodass die Elektrizität durch das Multimeter hindurchfließt (Abbildung 6 links). Da es sich bei der elektrischen Spannung um einen Potentialunterschied zwischen zwei Punkten eines Leiterabschnittes handelt, kann der interessierende Leiterabschnitt gemessen werden, ohne den Stromkreis zu öffnen (Abbildung 6 rechts).

Abbildung 6 links: Messung der Stromstärke, rechts: Messung der Spannung.



Die folgenden inhaltsbezogenen Kompetenzen aus dem Bildungsplan 2016 für die Sekundarstufe I in Baden-Württemberg spiegeln diese Grundlagen wider:

„Die Schülerinnen und Schüler können

- (1) grundlegende Bauteile eines elektrischen *Stromkreises* benennen und ihre Funktion beschreiben (unter anderem *Schaltsymbole*) [...]
- (2) qualitativ beschreiben, dass elektrische Ströme einen Antrieb beziehungsweise eine Ursache benötigen und durch Widerstände in ihrer Stärke beeinflusst werden (*Stromstärke*, *Potential*, *Spannung*, *Widerstand*, *Ladung*),
- (3) den elektrischen *Stromkreis* und grundlegende Vorgänge darin mithilfe von Modellen erklären,
- (4) den Aufbau eines *Stromkreises* unter Vorgabe einer *Schaltskizze* durchführen sowie *Stromkreise* in Form von *Schaltskizzen* darstellen,
- (5) *Stromstärke* und *Spannung* messen,
- (6) in einfachen *Reihenschaltungen* und *Parallelschaltungen*, Gesetzmäßigkeiten für die *Stromstärke* und die *Spannung* beschreiben (Maschenregel, Knotenregel)“ (MKJS, 2015a).

Außerdem wird im Rahmen der Leitgedanken zum Kompetenzerwerb Folgendes formuliert: „Kursiv geschriebene Fachbegriffe in den inhaltsbezogenen Kompetenzen (zum Beispiel Energie) sind im Unterricht verbindlich mit dem Ziel einzusetzen, dass die Schülerinnen und Schüler diese

- in unterschiedlichen Kontexten ohne zusätzliche Erläuterung verstehen und anwenden können,
- im eigenen Wortschatz als Fachsprache aktiv benutzen können,
- mit eigenen Worten korrekt beschreiben können.

Fachbegriffe, die in den Standards nicht kursiv gesetzt sind, werden verwendet, um die Kompetenzbeschreibung für die Lehrkräfte fachlich präzise und prägnant formulieren zu können. Die Schülerinnen und Schüler müssen über diese Fachbegriffe nicht verfügen können“. (MKJS, 2015a)

Neben der Berücksichtigung der Kompetenzen wurden außerdem die Inhalte mit den Lehrwerken abgeglichen, damit die Intervention an den Unterricht und die Schulcurricula anschließt. In allen betrachteten Büchern erfolgt nach einer Einheit zum Magnetismus der Einstieg in die Elektrizitätslehre mit dem einfachen elektrischen Stromkreis. Es erscheint vorteilhaft, sich zunächst mit den Stromkreisen zu beschäftigen und erst danach mit der Elektrostatik zu beginnen, da die Schüler:innen durch die Versuche leicht begeistert werden können. Üblicherweise werden auch die bedeutendsten Schaltzeichen und die Anwendung für den einfachen Stromkreis gezeigt. In allen Büchern wird auch der Aufbau der Glühlampe skizziert. Anschließend werden, mit unterschiedlichem Detaillierungsgrad, leitende und nichtleitende Materialien betrachtet. Danach wird eine Analogie hergestellt. In den betrachteten Büchern wird dazu immer der Wasserstromkreis herangezogen. Mit dieser Analogie werden die Größen *elektrische Stromstärke I*, *Spannung U* und *Widerstand R* eingeführt. Mithilfe eines einfachen Atommodells werden schließlich die Ladungen erklärt (Burzin et al., 2016; Carmesin et al., 2016; Kilian, 2017). Für eine kritische Auseinandersetzung mit verschiedenen Schulbüchern wird an dieser Stelle auf Wilhelm und Vairo Nunes (2019) verwiesen.

Ein großes Problem stellen dabei auch die inzwischen ausführlich dokumentierten Schüler:innenvorstellungen dar, die das Konzeptverständnis erschweren. In der physikdidaktischen Forschung wird unter dem Begriff *Schüler:innenvorstellungen* das bei Schüler:innen vorliegende Verständnis physikalischer Konzepte zusammengefasst (Viering et al., 2010). Um sie zu ermitteln,

werden falsche Testantworten und Aussagen aus Interviews betrachtet. Die für die Intervention relevanten Schüler:innenvorstellungen sollen im Folgenden kurz vorgestellt werden.

- Notwendigkeit des geschlossenen Stromkreises nicht bewusst:
Schüler:innen wissen zwar in der Theorie, dass Stromkreise geschlossen sein müssen, damit Elektrizität fließt. Jedoch zeichnen sie in Schaltskizzen nur einen Leiter vom Elektrizitätsantrieb zum Nutzer ein (Wilhelm & Hopf, 2018).
- Stromverbrauchsvorstellung:
Schüler:innen glauben, dass in Bauteilen Strom verbraucht wird, sodass bei Reihenschaltungen die Stromstärke bei Bauteilen nahe der Batterie größer ist (Schecker & Duit, 2018, S. 4–9).
- Konstantstromquelle:
Der Elektrizitätsantrieb liefert eine konstante Stromstärke. Daher glauben Schüler:innen, dass in Reihenschaltungen mit einer oder mehreren Glühbirnen in allen Stromkreisen die Stromstärke identisch ist (Rhöneck, 1986).
- Sequenzielles Denken:
Die Stromstärke spaltet sich an Knotenpunkten immer in gleich große Teile auf (Closset, 1984, S. 21). Damit verknüpft ist die Vorstellung, dass die Stromstärke bei Vergrößerung des Widerstands kleiner wird, da ein größerer Widerstand mehr Strom verbraucht (Urban-Woldron & Hopf, 2012).
- Mangelnde Unterscheidung zwischen Spannung und Stromstärke:
Die Spannung wird von vielen Schüler:innen nur als Beiwerk oder Eigenschaft der elektrischen Stromstärke betrachtet – oder sogar als Rechengröße. Es wird davon ausgegangen, dass dies auf eine mangelnde Unterscheidung zwischen Strom, Spannung und elektrischer Energie zurückzuführen ist (Rhöneck, 1986).
- Mangelnde Unterscheidung zwischen Reihen– und Parallelschaltung:
Befinden sich im Stromkreis mehr als zwei Nutzer, können Schüler:innen nicht mehr zwischen Parallel- und Reihenschaltungen unterscheiden (McDermott & Shaffer, 1992).
Die Symbole in Schaltplänen stellen für Schüler:innen ein großes Problem dar. Sie können in den Symbolen nicht die eigentlichen Bauteile wiedererkennen (Schecker & Duit, 2018).
- Konzeptionelle Schwierigkeiten mit der Spannung:
Schüler:innen gehen davon aus, dass an den beiden offenen Enden keine Spannung anliegt, wenn ein Leiter durchtrennt oder ein Schalter geöffnet wird (Burde, 2018).

Diesen Fehlvorstellungen soll in der Intervention durch passende Versuche und Aufgaben sowie sprachlich sensible Informationstexte entgegengewirkt werden. Beispielsweise können mithilfe von Multimetern die Richtung der Stromstärke sowie ihre Veränderung beim Einbringen mehrerer Nutzer beobachtet werden. Die angestrebte Wissensbasis wurde mit Lehrkräften besprochen, um an Zielgruppe und Curriculum angepasst zu sein (Prenzel et al., 2002; Walpuski & Ropohl, 2014), und umfasst die Grundgrößen der Elektrizitätslehre gemäß dem Bildungsplan Physik (MKJS, 2015a). Es wurde geplant, dass die Intervention am Anfang der Unterrichtseinheit zur Elektrizitätslehre steht, um so die Einbindung in den Unterricht zu vereinfachen (Annetta et al., 2013).

4.1.2 Das Serious Game Lights Out

Das Serious Game *Lights Out* wurde in der Engine Godot (Godot, 2023) entwickelt, und alle Animationen sowie Grafiken wurden in GIMP (GIMP, 2023) und Aseprite (Capello, 2023) erstellt.

Einerseits sollte das Spiel zeitlos gestaltet sein und nicht in Konkurrenz mit aktuellen AAA¹⁴-Spielen stehen, andererseits favorisiert die Zielgruppe in den Klassen 7 und 8 Spiele wie *Minecraft*, *Among Us* und andere (Feierabend et al., 2020). Daher wurde für das Spiel das Pixel-Art-Design gewählt. Ein weiterer Vorteil liegt in der kognitiven Entlastung: Durch die weniger visuell beladenen Grafiken kann der Lerninhalt stärker in den Fokus rücken. Des Weiteren wurde die Erzählung auf wenige Sequenzen und Texte beschränkt, sodass die Lerninhalte des Spiels im Vordergrund stehen. Durch die passende narrative Rahmung als Alltagsproblem (Bellotti et al., 2012; Cheng et al., 2014; Couceiro et al., 2013) – ein Haus soll nach einem Stromausfall durch geeignetes Kombinieren von Lampen und Batterien beleuchtet werden – soll zudem bei Mädchen das Interesse für die Elektrizitätslehre erhöht werden (Bressler et al., 2019; Dopatka et al., 2018). Durch das Serious Game kann auch die Angst vor Versuchen reduziert werden, da diese nur virtuell durchgeführt werden und so beim Experimentieren kein Versuchsmaterial zerstört werden kann (M. Schwarzer & Tschauko, 2010).

Das Spiel ist in vier aufeinander aufbauende Kapitel untergliedert, in denen die Lernenden zunehmend komplexere Themen der Elektrizitätslehre erarbeiten. Nachdem das Spiel gestartet wurde, erstellen die Schüler:innen einen persönlichen Avatar und tragen ihren Code als Namen für den Avatar ein (Abbildung 7).

Abbildung 7 Erstellung des Avatars.



Anschließend beginnt die Erzählung des Spiels, der Avatar findet sich in der Situation wieder, dass aufgrund eines Stromausfalles die Lichter ausgehen. Um sich im Haus zurechtzufinden, wird eine Beleuchtung benötigt, da auch die zunächst vorhandene Taschenlampe nur über begrenzte Kapazitäten verfügt (siehe Abbildung 8).

¹⁴ AAA bezeichnet die Blockbuster unter den Spielen, die große Budgets, riesige Produktionsteams, bekannte Publisher und entsprechend hohe Erwartungen an den Markterfolg kennzeichnen.

Abbildung 8 Im Haus können unterschiedliche Räume angesteuert werden.



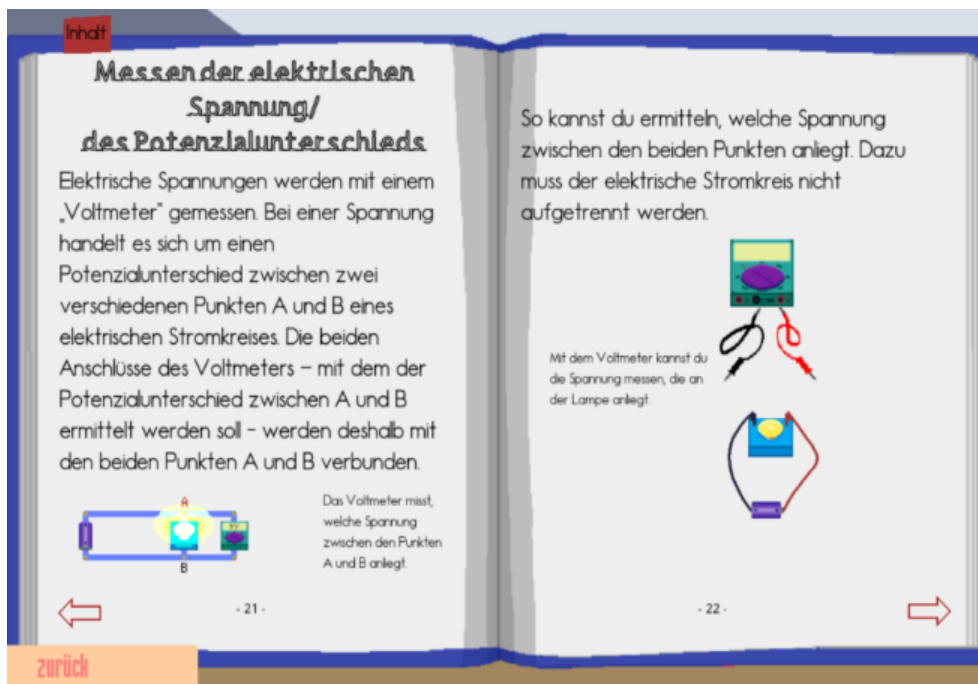
Das erste Kapitel beginnt mit der Einführung des Elektrizitätsbegriffs als Umschreibung für die elektrische Stromstärke und erarbeitet die Notwendigkeit des geschlossenen elektrischen Stromkreises sowie seiner Schaltzeichen. Im zweiten Kapitel wird der Begriff der elektrischen Stromstärke eingeführt, ohne auf das Modell elektrischer Ladungen oder Analogien zu Wasser einzugehen. Somit können Lehrkräfte in ihrem Unterricht das Modell ihrer Wahl verwenden. Im dritten Kapitel wird die elektrische Spannung über den Potentialunterschied erklärt. Eine Differenzierungsaufgabe für besonders schnelle Schüler:innen bildet das vierte Kapitel, in dem komplexere Schaltungen mit den Kirchhoffschen Gesetzen thematisiert werden. Eine Übersicht über die Themen der einzelnen Kapitel, die Lernziele und den Bezug zum Bildungsplan befindet sich im Anhang A.

Das Haus hat drei Stockwerke, in denen verschiedene Räume über die Schnittansicht des Hauses betreten werden können. Hier befinden sich auch ein Fortschrittsbalken und die Anzeige über gelöste und noch zu lösende Aufgaben (Abbildung 9). Im Wohnzimmer kann in einem Buch das benötigte Wissen zur Elektrizitätslehre nachgeschlagen werden (Abbildung 10). Die Texte sind möglichst einfach geschrieben und werden durch passende Animationen unterstützt. Die Schüler haben die Wahl, ob sie das Buch als Hilfestellung während des Spiels heranziehen oder im Voraus alles genau durchlesen. In den OECD-Studien wurde eine geschlechtsspezifische Präferenz beim Vorgehen ermittelt (OECD, 2016).

Abbildung 9 Lösen der Aufgaben im Kinderzimmer.



Abbildung 10 Buch zur Elektrizitätslehre.



Im ersten Stock befindet sich das Kinderzimmer. Dort werden die Aufgaben zu den jeweiligen Themen bearbeitet. Der Link zu den passenden Seiten im Buch erscheint als Hinweis, wenn die Aufgabe nicht gelöst werden kann. Vollständig fehlerfrei gelöste Aufgaben werden mit drei Sternen belohnt. Wenn Schaltkreise analysiert werden müssen, kann die Antwort auf dem Dachboden getestet werden, bevor sie abgegeben wird. Außerdem besteht die Möglichkeit, die jeweilige Aufgabe auf den Dachboden zu nehmen und die Lösung dort zu erarbeiten.

Auf dem Dachboden können Versuche zur Elektrizitätslehre durchgeführt werden, indem Lampen, Batterien und Leiter zu verschiedenen Schaltungen kombiniert werden (Abbildung 11). Im weiteren Spielverlauf werden noch Messgeräte, Amperemeter und Voltmeter erworben.

Abbildung 11 Experimente auf dem Dachboden.



Im letzten Raum, der Küche, können die genauen Lernfortschritte in Form von Badges eingesehen werden (Abbildung 12). *Badge* ist eine im Kontext des digitalen Spielens und Lernens gebräuchliche Bezeichnung für Abzeichen, die die erworbenen Kenntnisse bestätigen. Beispielsweise erhalten Spielende das Badge *Reihenschaltung*, wenn sie alle Aufgaben und Versuche in Zusammenhang mit der Reihenschaltung erfolgreich bearbeitet haben.

Abbildung 12 Badges für erfolgreich bearbeitete Inhalte.



Im Serious Game sollen die Phasen, in denen sich die Schüler:innen mit dem Erlernen der Spielsteuerung beschäftigen, möglichst kurz sein. Um dies zu erreichen, wurden beim Programmieren

die zwei folgenden Punkte umgesetzt: Zum einen wird die Steuerung in einem kurzen Intro erläutert, das jederzeit über das Escape-Menü erneut aufgerufen werden kann und möglichst intuitiv ist, zum anderen wurden kurze Tutorial-Videos direkt im Spielgeschehen untergebracht (Gee, 2007).

Interaktivität wird durch ein flexibles Vorgehen beim Bearbeiten der Aufgaben geschaffen. Den Lernenden wird möglichst viel Handlungsfreiheit gewährt, indem sie wählen können, ob sie zuerst Fachtexte lesen, Aufgaben lösen oder Versuche durchführen. Bevor die Lernenden zum nächsten Kapitel übergehen können, müssen alle Aufgaben und Versuche des letzten Kapitels abgeschlossen sein (Kiili, 2005; Ravysse et al., 2017). Die Interaktivität soll außerdem dazu beitragen, dass Schüler:innen Flow erleben. Dazu wurden die folgenden Komponenten des Flow-Erlebens (Csikszentmihalyi, 2014) berücksichtigt:

- Klare Ziele und Teilziele: Die übergeordnete Problemstellung besteht darin, eine Beleuchtung für das Haus zu erstellen. Im Zuge der Lösungssuche können sich Lernende auch mit anderen Wirkungen der Elektrizität beschäftigen. Der Schwierigkeitsgrad der Aufgaben und Kapitel nimmt im Spielverlauf zu, und Items werden erst nach und nach freigeschaltet, sodass nicht von Anfang an zu viele Möglichkeiten bestehen. Gelöste Aufgaben werden durch das Freischalten neuer Items belohnt. Wie viele Aufgaben gelöst wurden und wie viele noch übrigbleiben, wird pro Kapitel und für das Gesamtspiel jederzeit transparent gemacht und im Haus angezeigt. Für den Inhaltsbereich relevante Übungen können den Lernenden Fähigkeiten vermitteln, auf die sie sich konzentrieren können (Jones, 1998).
- Passung von Fähigkeit und Anforderung: Durch die Unterteilung in Module mit einzelnen Lerninhalten wird der Schwierigkeitsgrad langsam gesteigert, ohne dass Vorwissen benötigt wird. Die Lernenden können sich Wissen aneignen, entsprechende Versuche durchführen und ihre Kenntnisse in einem Test überprüfen. Informationstexte können, müssen aber nicht abgerufen werden, wenn sich jemand lieber auf die Aufgaben oder Versuche konzentriert. Die Aufgaben in den Kapiteln können in beliebiger Reihenfolge abgeschlossen werden. Auch ist es möglich, bereits gelöste Aufgaben erneut zu bearbeiten.
- Klares, informatives Feedback: Zu jedem Zeitpunkt können Hilfestellungen und Tipps abgerufen werden. Durch das Feedback erhalten die Lernenden Rückmeldung über den Erfolg bei den Aufgaben sowie zu ihrem derzeitigen Lernstand.
- Eindeutigkeit der Handlungsstruktur: Jones (1998) empfiehlt hierfür eine nahtlose Integration von „tools, tasks and presentation of information“. Zu jedem Zeitpunkt kann zwischen den Informations-, Übungs- und Testgelegenheiten gewechselt werden. Zum Beispiel ist es möglich, einen Versuch zum Kurzschluss durchzuführen, darüber zu lesen und entsprechende Fragen zu beantworten.

Wie in Abschnitt 2.3 dargestellt, wird vermutet, dass die Elemente des Flow-Erlebens auch in Verbindung mit den Quellen der Selbstwirksamkeitserwartung stehen. Daher wurden bei der Gestaltung des Spiels die vier Quellen der Selbstwirksamkeitserwartung (siehe Tabelle 3) nach Bandura (1977) berücksichtigt (Perlwitz & Stemmann, 2022; Perlwitz et al., 2022). So werden als erste Quelle der Selbstwirksamkeitserwartung Erfolgserlebnisse erreicht, indem in einer risikofreien Umgebung Aufgaben in steigendem Schwierigkeitsgrad bearbeitet werden (Bergey et al., 2015; Jarvis & Freitas, 2009). Zusätzlich gibt ein Fortschrittsbalken Aufschluss über bereits erreichte Ziele und gelöste Aufgaben werden durch Game-Elemente, wie Sternen, Badges und freigeschalteten Gegenständen belohnt. Die stellvertretende Erfahrung, die zweite Quelle, soll durch einen individuell anpassbaren Avatar ermöglicht werden. Zusätzlich können sich die Schüler so stärker mit dem Spielgeschehen identifizieren. Der Avatar gibt zudem positives Feedback und Hilfestellungen und fungiert damit als dritte Quelle der verbalen Rückmeldung. Zu den Interaktionsmöglichkeiten im Spiel gehören außerdem

eine Chat-Schnittstelle, worüber mit der Lehrkraft und den Mitschüler:innen die Probleme diskutiert werden können. Auf ein kompetitives Design, z. B. mit Ranglisten oder Ähnlichem, wurde verzichtet, da soziale Bezugsnormierungen die Unsicherheit vergrößern können. Insbesondere bei Schüler:innen mit geringer Selbstwirksamkeitserwartung könnte dadurch die Lernmotivation gehemmt werden, was sich negativ auf die vierte Quelle, die emotionale Erregung durch das Spiel, auswirken würde (Dweck, 1986; Werbach et al., 2012).

Ergebnisse der Pretests mit dem Serious Game

Das Spiel wurde zunächst von einer Expert:innengruppe getestet. Diese umfasste Lehramtsstudierende und Lehrkräfte der Fächer Technik und Physik, Personen aus dem Bereich Software-Entwicklung und -Testing, Physiker:innen, E-Sportler:innen und Personen, die sich nie mit Computerspielen beschäftigen und weder physikalische noch technische Vorbildung hatten. In einem Formular (Anhang B) konnten die Teilnehmenden mögliche Fehler und im Verlauf des Spiels aufgetretene Schwierigkeiten eintragen. Die Rückmeldungen gaben nicht nur Aufschluss über die Schwierigkeiten des Spiels, die mögliche Spieldauer sowie vorhandene Bugs, sondern auch darüber, welche Elemente die Spielenden als Flow-fördernd oder -hemmend empfanden. Die Expert:innen benötigten durchschnittlich zwei Zeitstunden, um das Spielziel zu erreichen. Dabei waren Personen mit fundiertem Wissen in Elektrotechnik deutlich schneller als Personen, die weder Kenntnisse der Elektrizitätslehre noch Erfahrungen mit Computerspielen hatten.

Das Spiel wurde anschließend überarbeitet und ohne Kontrollgruppe mit 79 Schüler:innen aus einer 7. Klasse und zwei 8. Klassen eines Gymnasiums getestet. Der Fokus lag zunächst darauf, die Spieldauer zu verifizieren und die Schwierigkeiten oder Frustrationen zu identifizieren, die bei der Bearbeitung auftreten können. Zusätzlich füllten die Teilnehmenden die Fragebögen mit den Skalen zur Selbstwirksamkeitserwartung, den Selbstkonzepten und dem Fachwissen aus, um die Dauer der Bearbeitung abschätzen zu können.

Des Weiteren wurden, ähnlich wie zuvor bei der Expert:innen-Testung, Formulare ausgegeben, auf denen Fehlfunktionen des Spiels, aber auch besonders hilfreiche und besonders störende Aspekte eingetragen werden konnten. Des Weiteren gab es auf den genannten Formularen für die Schüler:innen die Möglichkeit, Begriffe zu nennen, die ihnen im Spiel oder bei den Fragebögen Schwierigkeiten bereitet haben.

Insgesamt dauerte die Durchführung vier Schulstunden. Zum einen benötigte das Ausfüllen der Fragebögen viel Zeit, zum anderen traten noch Fehler im Spiel auf, sodass einige Schüler:innen gezwungen waren, das Spiel neu zu beginnen.

Daher wurde entschieden, in der Intervention nur drei der vier Kapitel einzusetzen und das vierte Kapitel als Differenzierung für schnellere Schüler:innen zu verwenden.

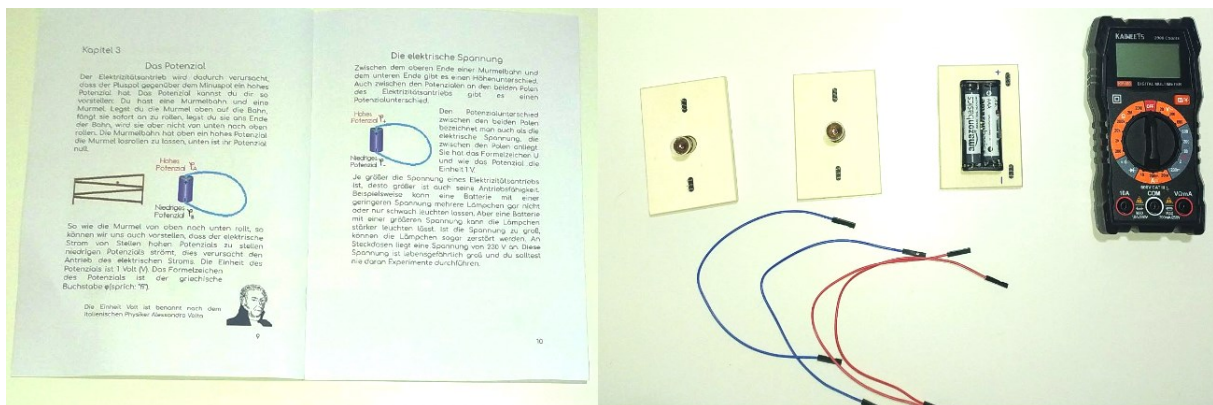
Mit dieser ersten Testung konnten insgesamt 29 vollständige Datensätze gewonnen werden. Die geringe Zahl ist zum einen darauf zurückzuführen, dass die Schüler:innen pandemiebedingt nicht über die gesamten vier Stunden anwesend waren. Zum anderen blieben viele Datensätze auch aufgrund von verlorenen und vergessenen Codes, verlorenen Spielständen und Schwierigkeiten mit der Internetverbindung unvollständig. Die Papierbögen für die Rückmeldungen zum Spiel wurden teilweise gewissenhaft ausgefüllt. Dabei wurde mehrheitlich positiv hervorgehoben, dass es einen individuellen Avatar gibt und ein Buch, in dem alles nachzulesen ist. Auch das Pixel-Art-Design wurde gelobt. Mehrere Schüler:innen sahen es auch als Vorteil, sich ohne Druck mit Inhalten beschäftigen zu können. Negative Anmerkungen betrafen vor allem den Umstand, dass das Spiel gelegentlich noch abgestürzt ist und dass die Flow-Fragen die Schüler:innen in ihrer Konzentration unterbrechen. Letzteres kann als Hinweis auf ein mögliches Flow-Erleben gedeutet werden.

Die Fehler im Spiel, die die Kinder gemeldet hatten, wurden behoben und das Spiel wurde erneut von verschiedenen Expert:innen getestet. Anschließend fand die erste vollständige Pilotierung der Materialien, dieses Mal mit Kontrollgruppe, statt.

4.1.3 Materialien der Kontrollgruppe

Für die Kontrollgruppe wurden Interventionsmaterialien entwickelt, die einen möglichst fairen Vergleich mit der Experimentalgruppe erlaubten (Abbildung 13). Die Lernziele, Inhalte, Aufgaben und Versuche sollten mit denen aus dem Serious Game identisch sein. Hierfür wurde ein Stecksystem entwickelt, mit dem die Schüler:innen dieselben Schaltungen wie im Serious Game aufbauen konnten. Das Stecksystem bestand aus Nutzern und Elektrizitätsantrieben. Zu diesem Zweck wurden Glühbirnenfassungen mit Glühbirnen bzw. Batteriehalterungen auf kleinen Holzplatten montiert und mit Pins versehen. Mit Steckbrücken-Kabeln konnten diese Pins verbunden und so verschiedene Reihen- und Parallelschaltungen, ähnlich wie im Serious Game, realisiert werden. Einfache Multimeter wurden mit Messleitungen mit Krokodilklemmen ausgestattet, sodass an den Pins die elektrische Stromstärke und die Spannung gemessen werden konnten.

Abbildung 13 Materialien der Kontrollgruppe.



Jede:r Schüler:in der Kontrollgruppe erhielt einen Kasten, der vier Lampen, zwei Batterien, ein Multimeter, zehn Steckbrücken-Kabel und zwei Messleitungen enthielt.

Zusätzlich befand sich in jedem Kasten ein Elektrizitätslehre-Heft, in dem die Informationstexte und einzelne Bilder der Animationen aus dem Buch des Serious Games abgedruckt waren. Ergänzend wurden Bilder der Schaltungen mit dem Stecksystem eingefügt. Die Aufgaben und Versuche wurden auf Papier ausgegeben und entsprachen weitgehend denen des Serious Games. Lediglich an Stellen, an denen konkrete elektrische Stromstärken oder Spannungen abgefragt wurden, erfolgte eine Anpassung in den Aufgabenstellungen, um den Werten der realen Lampen und Batterien zu entsprechen. Während der Intervention konnten die Schüler:innen ihre Ergebnisse mit Musterlösungen abgleichen, die in einfacher Ausfertigung auf dem Lehrkraft-Pult auslagen.

4.1.4 Fachwissenstest

Der Fachwissenstest diente dazu, Transparenz über das vorhandene sowie gewonnene Fachwissen im Bereich der Elektrizitätslehre zu schaffen (J. Rost, 2004, S. 43). Daher sollten möglichst unterschiedlich schwierige Aufgaben das Fachwissen der Proband:innen repräsentieren. Ziel war es, ausschließlich zu prüfen, inwieweit die Konzepte und Prinzipien vollständig erfasst wurden. Es wurde dabei keine Differenzierung in Leistungsbereiche bzw. nach der Geschwindigkeit bei der Textbearbeitung angestrebt (Bühner, 2011).

Fachwissen gehört zu den vier naturwissenschaftlichen Kompetenzbereichen der Bildungsstandards (Weinert, 2001). Zum Fachwissen in Physik zählen Fakten, Begriffe, Prinzipien, Modelle und Gesetzmäßigkeiten sowie Basiskonzepte zur Lösung konkreter Probleme. Die Kompetenzbereiche werden in drei Anforderungsbereiche unterteilt, die in der Tabelle 7 zu sehen sind.

Tabelle 7 Anforderungsbereiche des Fachwissens (KMK, 2004).

Anforderungsbereiche Fachwissen		
I	II	III
<i>Wissen wiedergeben</i>	<i>Wissen anwenden</i>	<i>Wissen transferieren und verknüpfen</i>
Fakten und einfache physikalische Sachverhalte reproduzieren.	Physikalisches Wissen in einfachen Kontexten anwenden, einfache Sachverhalte identifizieren und nutzen, Analogien benennen.	Wissen auf teilweise unbekannte Kontexte anwenden, geeignete Sachverhalte auswählen.

Die Inhalte des Fachwissenstests entsprechen gemäß dem Kompetenzmodell der Kultusministerkonferenz (KMK, 2004) für Fachwissen den Anforderungsbereichen I und II (Hattie, 1999). Anforderungsbereich I umfasst die Kompetenz, Wissen aus einem abgesteckten Bereich wiederzugeben und in einem vergleichbaren Zusammenhang direkt anzuwenden. Ein Beispiel dafür ist die Aufgabe, in der ein Schaltsymbol der Elektrizitätslehre benannt werden soll. Im Anforderungsbereich II wird die Kompetenz erwartet, mehrere überschaubare Bereiche miteinander in Zusammenhang zu bringen und zu erläutern, zu interpretieren und bestimmte Informationen zu identifizieren. Eine Aufgabe für diesen Anforderungsbereich ist die Ermittlung der Stromstärke in bestimmten Leitern in einer Schaltung mit Verzweigungen bei gegebenen Widerständen und gegebener Gesamtspannung.

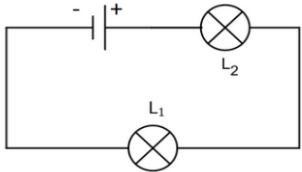
Nicht nur aus Gründen der Zeitökonomie und der Objektivität sind geschlossene Antwortformate empfehlenswert, zwischen denen gewählt werden kann (Bortz & Döring, 2016). Ein weiterer Grund besteht darin, dass die Fachsprache im naturwissenschaftlichen Unterricht Satzbau und Sprachstil nicht der Alltagssprache entsprechen (Leisen, 2005). Dies führt dazu, dass die sprachlichen Fähigkeiten der Schüler:innen im naturwissenschaftlichen Unterricht eine mögliche Varianz mit sich bringen (O'Reilly & McNamara, 2007). Daher scheint das Multiple-Choice-Format für den Fachwissenstest mit drei Distraktoren vorteilhaft (Rodriguez, 2005). Bei der Erstellung ist auch die Art der Distraktoren von Bedeutung: Sind sie zu leicht, wird die Antwort offensichtlich, sind sie zu schwer, können selbst Personen mit viel Wissen die richtige Antwort nicht mehr erkennen (Bühner, 2011). Eine Kategorie *Ich weiß nicht* wurde bewusst nicht erstellt, da so einerseits die Anzahl korrekter Antworten auf bis zu 15 % gesteigert werden kann (Mondak & Davis, 2001), andererseits die Ratewahrscheinlichkeit verringert wird (ebd.). Die Proband:innen wurden angehalten, bei Unsicherheit die wahrscheinlichste Antwort zu wählen. Geeignete Distraktoren können darüber hinaus Aufschluss über die vorhandenen Vorstellungen der Schüler:innen, insbesondere deren Fehlvorstellungen, geben. Die Fehlvorstellungen in der Elektrizitätslehre wurden über Jahrzehnte aus falschen Testantworten und Interviews mit Schüler:innen ermittelt¹⁵. Schüler:innen gehen nicht mit den grundsätzlich gleichen Fehlvorstellungen an Aufgaben heran. Sie haben individuelle Lösungsstrategien, die auch bei falschen Vorstellungen zur

¹⁵ Einige der Messinstrumente: McDermott und Shaffer (1992); Rhöneck (1986).

richtigen Antwort führen. Solche falschen Antworten bei richtiger Begründung oder richtigen Antworten, aber falsche Begründung werden als Fehler erster und zweiter Art bezeichnet. Das Testinstrument von Urban-Woldron und Hopf (2012) berücksichtigt dies durch ein zweistufiges Antwortformate, mit diesem wird in der ersten Frage das Fachwissen abgefragt und in einer Folgefrage eine Begründung für die Antwort der ersten Frage. Neben der Berücksichtigung von mehrstufigen Antworten mit jeweiligen Distraktoren, ist ein weiterer Vorteil des Tests der Einbezug verschiedener Unterrichtszugänge. Dies ermöglicht einen fairen Vergleich der verschiedenen konkurrierenden Zugänge zu den Themen der Elektrizitätslehre.

Für den Fachwissenstest dieser Arbeit (siehe Anhang D) diente das Testinstrument von Urban-Woldron und Hopf (2012) als Grundlage, und viele Aufgaben wurden daraus entnommen. Bei der ersten Betrachtung fielen sprachliche Probleme des Instruments auf, bspw. eine uneinheitliche Verwendung der Begriffe *Glühlampe*, *Glühbirne* und *Lämpchen*. Daher wurden in einem ersten Schritt die Begriffe angeglichen und die Aufgaben sprachlich vereinfacht (Jonkisz et al., 2012). Einzelne Aufgaben wurden dem Testinstrument von Sokoloff (1996) entnommen und aus der Dissertation von Burde solche Aufgaben, bei denen die Spannung abgefragt wird (Burde, 2018). Beide Instrumente wurden von den Ersteller:innen hinsichtlich Inhaltsvalidität und Expertenvalidität geprüft und es wurden gute Reliabilitäten (Cronbachs $\alpha = .84 - 85$) angegeben (Urban-Woldron & Hopf, 2012; Burde 2018). Außerdem wurden Aufgaben, die nicht den Lerninhalt des Spiels betreffen, sowie solche mit zu komplexen Schaltungen oder mit unterschiedlichen Widerständen verändert oder gestrichen, sodass in den Aufgaben nur Glühlampen gleichen Widerstands vorkamen. Ein Beispielitem für eine Aufgabe nach Urban-Woldron und Hopf (2012) ist in Abbildung 14 zu sehen.

Abbildung 14 Beispielaufgabe aus dem Fachwissenstest nach Urban-Woldron Hopf (2012).

Betrachte die Abbildung. Wie hell werden die Glühlampen L_1 und L_2 leuchten?	
<input type="checkbox"/> Beide Glühlampen leuchten gleich hell. <input type="checkbox"/> L_2 leuchtet. L_1 leuchtet nicht. <input type="checkbox"/> Beide Glühlampen leuchten. L_2 leuchtet heller als L_1 . <input type="checkbox"/> L_1 leuchtet. L_2 leuchtet nicht.	
Wie begründest du deine Entscheidung?	
<input type="checkbox"/> Der elektrische Strom ist überall im Stromkreis gleich. <input type="checkbox"/> L_2 verbraucht einen Teil des Stroms. Es ist daher weniger Strom für L_1 übrig. <input type="checkbox"/> L_2 ist näher bei der Batterie. Daher bekommt sie mehr Strom. <input type="checkbox"/> Der Strom wird gleichmäßig auf beide Glühlampen aufgeteilt.	

Die vollständige Auflistung der Items sowie ihre Zuordnung zu Kompetenzen und Fehlvorstellungen befinden sich im Anhang D.

Die Auswertung des Multiple-Choice-Formats erfolgte dichotom (Cohen, 1983). Somit konnte die Schüler:innenvorstellung entweder vorhanden (Wert 1) oder nicht vorhanden sein (Wert 0). Der Summenscore berechnete (Prenzel et al., 2002) sich aus dem prozentualen Anteil richtig beantworteter Fragen. Der Fachwissenstest wurde zur Überprüfung der Schwierigkeit und der

Inhaltsvalidität einer Expert:innengruppe (Moosbrugger & Kelava, 2020, S. 15) aus Lehrkräften des Fachbereichs Physik (N = 9) vorgelegt. Die Inhaltsvalidität zeigt auf, inwieweit die Testinhalte das relevante Merkmal repräsentativ erfassen und für die Messung geeignet sind (Eid & Schmidt, 2014, S. 56). Insgesamt wurden zwölf einstufige und 14 zweistufige Items auf zwei Testhefte (A und B) aufgeteilt und zunächst ohne die Interventionsmaterialien zur Elektrizitätslehre pilotiert.

Um herauszufinden, ob beide Tests in Inhalt und Schwierigkeit gleichwertig sind und für dieselbe Gruppe von Testteilnehmern ähnliche Ergebnisse erzielen, wurde die Paralleltest-Reliabilität ermittelt (Moosbrugger & Kelava, 2020, S. 125).¹⁶

Die Paralleltest-Reliabilität wurde vor der Pilotierung bei Schüler:innen aus zwei 8. Klassen eines Gymnasiums erhoben. Beide Klassen hatten das Thema Elektrizitätslehre abgeschlossen und erarbeiteten seit zwei Wochen das Thema *Mechanik*, sodass keine Veränderung des Lernstandes in der Elektrizitätslehre erwartet wurde. Zwischen den beiden Testungen lagen zwei Wochen Osterferien, sodass Erinnerungseffekte nahezu ausgeschlossen werden konnten. Aufgrund der COVID-Pandemie fehlte nach den Osterferien ein Großteil beider Klassen, sodass nur 32 Schüler:innen den Test vollständig bearbeiteten. Im Mittel erreichten die Schüler:innen im A-Test $M = .40$ ($SD = .15$) und im B-Test $M = .38$ ($SD = .17$) Punkte. Die Pearson-Korrelation der beiden Test-Varianten ergab mit $r = .6$ zwar ein zufriedenstellendes Ergebnis, allerdings ist die statistische Aussagekraft aufgrund des kleinen Datensatzes gering. Eine Wiederholung gestaltete sich während der Pandemie als schwierig. Es wird vermutet, dass die Reliabilität paralleler Tests robuster ist, wenn die Anzahl der Items im Test größer ist, da bei einer kleineren Anzahl von Elementen der Korrelationskoeffizient möglicherweise nicht groß genug ist (Brennan, 2001). Des Weiteren wurde der Fachwissenstest bei 25 Schüler:innen einer 9. Klasse erprobt. Erwartungsgemäß wurden die Fragen hier besser beantwortet mit Mittelwerten beim A-Fachwissenstest von $M = .56$ ($SD = .20$) und beim B-Fachwissenstest von $M = .53$ ($SD = .17$). Insgesamt zeigte sich anhand der Mittelwerte, dass der Test eher als schwierig einzustufen ist.

4.1.5 Erfassung des Flow-Erlebens

Die Messung des Flow-Erlebens stellt einen komplexen Prozess dar, und verschiedene Methoden sowie Maßnahmen können je nach Kontext und untersuchter Population zu unterschiedlichen Ergebnissen führen. Im Kontext der Flow-Messungen finden unter anderem folgende Methoden Anwendung: Beobachtungen, Interviews, Fragebögen, Verhaltensmessungen und physiologische Messungen. Aus Gründen der Objektivität und Zeitökonomie soll das Flow-Erleben mit einem Fragebogen erhoben werden. Da bei Selbstauskünften, die im Nachhinein erhoben wurden, sich Erinnerungsverzerrungen aufdecken ließen (Brewer et al., 1991), soll das Flow-Erleben während der Intervention ermittelt werden. Ein möglicher Nachteil besteht bei diesem Vorgehen darin, dass die Proband:innen durch das Beantworten der Fragen aus einem möglichen Flow-Zustand gerissen werden. Daher wurde ein Test-Instrument gewählt, das durch seinen geringen Umfang von zehn Items vorteilhaft erscheint: die Flow-Kurz-Skala (Rheinberg et al., 2019). Die Skala erfasst das Flow-Erleben mit seinen verschiedenen Komponenten und überzeugte durch eine hohe interne Konsistenz

¹⁶ Gelegentlich findet sich in der Literatur die Überprüfung der Güte der Reliabilität von Fragebögen durch Cronbachs α , woraus dann die Eindimensionalität des zugrundeliegenden Konstrukts gefolgert wird. Jedoch ist Eindimensionalität eine Annahme, um Cronbachs α anwenden zu können. Dem Fachwissenstest liegen sinngemäß mehrere Faktoren zugrunde, da unterschiedliches Wissen in unterschiedlicher Anforderung abgefragt wird. Daher ist die Prüfung mit Cronbachs α in diesem Fall nicht sinnvoll. Des Weiteren handelt es sich um dichotomisierte Items, deren Faktorisierung nur durch Überführung bspw. in tetrachorische Korrelationsmatrizen ermittelt werden kann, was jedoch erst ab Stichproben > 500 und bei weniger als 16 Items empfohlen wird (Y. Yang & Xia, 2019).

(Cronbachs α zwischen .80 und .90, Engeser & Rheinberg, 2008). Die Untersuchung der faktoriellen Validität ergab, dass die Items auf zwei Subfaktoren luden (ebd.). Des Weiteren wurde das Item in Settings mit Computerspielen eingesetzt (Rheinberg & Vollmeyer, 2003). Die Items sind in Tabelle 8 aufgeführt.

Tabelle 8 Items der Flow-Kurz-Skala nach Rheinberg et al. (2019).

Nr.	Item
FL1	Ich fühle mich optimal beansprucht. Ich bin ganz in meinem Element.
FL2	Meine Aktivitäten laufen flüssig und glatt.
FL3	Ich merke gar nicht, wie die Zeit vergeht.
FL4	Ich habe keine Mühe, mich zu konzentrieren.
FL5	Mein Kopf ist völlig klar.
FL6	Ich bin ganz vertieft in das, was ich gerade mache.
FL7	Die richtigen Gedanken kommen wie von selbst.
FL8	Ich weiß bei jedem Schritt, was ich zu tun habe.
FL9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.
FL10	Ich bin völlig selbstvergessen.

Der Fragebogen wurde in mehreren Studien validiert. Rheinberg und Vollmeyer (2003) nutzten den Fragebogen auch für Computerspiele (Cronbachs α von .8–.9). Die Items sollen während des Spiels in beliebiger Reihenfolge als Pop-up-Fenster auftreten.

Abbildung 15 Ein Pop-up-Fenster zur Abfrage des Flow-Erlebens im Serious Game.



Die Lernenden haben die Möglichkeit, die fünfstufige Likert-Skala (1 = *stimme nicht zu*, 5 = *stimme voll zu*) direkt oder später anzuklicken und ohne Unterbruch mit dem Spiel fortzufahren (siehe Abbildung 15). Auch bei der Kontrollgruppe wurde mit der Papier-und-Bleistift-Variante der Flow-Kurz-Skala der erlebte Flow erhoben. Dabei forderte die Testleitung die Teilnehmenden zu zufälligen Zeitpunkten dazu auf, je eine Frage zu beantworten.

Nach der ersten Testung des Serious Games wurde das Item FL1 textlich verändert und Item FL2 gestrichen. Die Schüler:innen beschrieben diese Items als schwer verständlich und konnten die Art der Antworten nicht einordnen, weshalb bei diesen eine große Varianz entstand. Die Ergebnisse der Reliabilitätsprüfung sind im Kapitel 5 aufgeführt.

4.1.6 Selbstwirksamkeitserwartung

Die Selbstwirksamkeitserwartung ist ein aufgabenspezifisches Konstrukt, da es sich auf die Ausführung bestimmter Aufgaben und Verhaltensweisen bezieht (Bandura, 1986). Daher sollten die Items auf die jeweilige Domäne zugeschnitten werden. Die in der Studie verwendeten fünf Items zur schulbezogenen Selbstwirksamkeitserwartung von Jerusalem und Satow (2009) wurden an das Fach Physik angepasst (siehe Tabelle 9, kursive Begriffe wurden ergänzt, um zum Fach Physik zu passen). Die Skala zur schulischen Selbstwirksamkeitserwartung ist ein etabliertes Instrument, für das eine akzeptable interne Konsistenz (*Cronbachs α = .69 - .72*) angegeben wurde. Die Werte wurden auf einer fünfstufigen Likert-Skala (1 = *stimme nicht zu*, 5 = *stimme voll zu*) jeweils vor (Selbstwirksamkeitserwartung(t_1)) und nach der Intervention (Selbstwirksamkeitserwartung(t_2)) erhoben.

Tabelle 9 Items zur Erfassung der physikbezogenen Selbstwirksamkeitserwartung (verändert nach Jerusalem & Satow, 2009).

Nr.	Item
SWE1	Ich kann auch die schwierigen Aufgaben <i>in Physik</i> lösen, wenn ich mich anstrenge.
SWE2	Wenn ich eine schwierige <i>Physik</i> aufgabe an der Tafel lösen soll, glaube ich, dass ich das schaffen werde.
SWE3	Selbst wenn ich mal längere Zeit krank sein sollte, kann ich immer noch gute Leistungen <i>in Physik</i> erzielen.
SWE4	Auch wenn meine <i>Physiklehrerin</i> oder mein <i>Physiklehrer</i> an meinen Fähigkeiten zweifelt, bin ich mir sicher, dass ich gute Leistungen erzielen kann.
SWE5	Ich bin mir sicher, dass ich auch dann noch meine gewünschten Leistungen <i>in Physik</i> erreichen kann, wenn ich mal eine schlechte Note bekommen habe.

4.1.7 Physikbezogenes und computerbezogenes Selbstkonzept

Das physikbezogene Selbstkonzept wurde mit Items (siehe Tabelle 10) aus der Dissertation von Weißnigk (2012) ermittelt, die für die Skalen eine hohe Varianzaufklärung (61,1% - 64%) und gute interne Konsistenz (*Cronbachs* $\alpha = .92 - .93$) angab. Das computerbezogene Selbstkonzept wurde mit den Items der Subskalen *Konativ* und *Kognitiv* (siehe Tabelle 11) aus dem Fragebogen von Janneck et al. (2012) erhoben. Diese Subskalen wurden gewählt, da diese Studie die bisherigen Erfahrungen (konativ) und der subjektiven Kenntnisstand (kognitiv) der Schüler:innen von Interesse war. Aufgrund von möglichen Verständnisschwierigkeiten der Subskala *Motivational* wurde auf diese verzichtet. Die Skala wurde bereits vielfach eingesetzt und zeichnet sich durch eine gute Validität (*CFI* > .95, *RMSA* < .05) und akzeptable bis sehr gute interne Konsistenz (*Cronbachs* $\alpha = .61 - .90$) Die Werte wurden vor der Intervention auf einer fünfstufigen Likert-Skala (1 = *stimme nicht zu*, 5 = *stimme voll zu*) gemessen.

Angesichts der Annahme, dass sich die Selbstkonzepte aufgrund ihrer stabilen Natur (siehe Kapitel 2.4) durch die Intervention nicht verändern, wird in dieser Arbeit eine einmalige Erhebung dieser Konzepte durchgeführt. Es wird vermutet, dass das computerbezogene Selbstkonzept ein Prädiktor des Flow-Erlebens in *Serious Games* ist [Klicken oder tippen Sie hier, um Text einzugeben.](#) und das physikbezogene Selbstkonzept ein Prädiktor für das Flow-Erleben bei der Auseinandersetzung mit physikalischen Inhalten.

Tabelle 10 Items zur Erfassung des physikbezogenen Selbstkonzept (verändert nach Weßnigg, 2012).

Nr.	Item
SKP 1	Ich traue mir einiges zu in Physik.
SKP 2	Das Fach Physik liegt mir nicht besonders.
SKP 3	In Physik weiß ich schon im Voraus, dass ich nichts verstehe.
SKP 4	Ich glaube, dass mich meine Mitschüler für wirklich gut halten in Physik.
SKP 5	Kein Mensch kann alles – ich habe einfach keine Begabung für Physik.
SKP 6	Wenn ich in Physik eine neue Aufgabe bearbeiten soll, bin ich sicher, dass ich sie schaffen werde.
SKP 7	Wenn eine Aufgabe in Physik kompliziert und schwierig erscheint, spornt mich das erst recht an, die Aufgabe zu lösen.

Tabelle 11 Items zur Erfassung des computerbezogenen Selbstkonzepts (Janneck et al., 2012).

Nr.	Item
SKC 1	Ich verfüge über viel praktische Erfahrung im Umgang mit Computern.
SKC 2	In meiner Freizeit beschäftige ich mich viel mit Computern (z. B. Computerspiele, Internet, Foto-, Videobearbeitung).
SKC 3	Ich halte mich im Umgang mit Computern für sehr kompetent.
SKC 4	Im Umgang mit Computern bin ich sicherer als der Durchschnitt.
SKC 5	Ich habe umfassende Computerkenntnisse.
SKC 6	Ich habe keine Scheu davor, neue Computerprogramme einfach auszuprobieren.
SKC 7	Eine neue Software probiere ich meist erst einmal aus.

Die Reliabilität der Skalen zur Erfassung des Flow-Erlebens, der Selbstwirksamkeitserwartung und des physikbezogenen und computerbezogenem Selbstkonzepts wurde mit McDonalds ω ermittelt (Brennan, 2001). Laut den Empfehlungen der Literatur sollte der Reliabilitätskoeffizient mindestens .7 betragen (Nunnally & Bernstein, 2008). Anderen Quellen zufolge sind allerdings bei kürzeren Tests auch kleinere Werte akzeptabel (Kopp & Lois, 2014; McCowan & McCowan, 1999; Schmitt, 1996). Die Ergebnisse sind in Kapitel 5.

4.1.8 Weitere Variablen

Zu den weiteren Variablen, die von den Teilnehmenden erhoben wurden, gehörte neben dem Alter das Geschlecht (männlich = 0, weiblich = 1). Wie bereits theoretisch hergeleitet, wird von geschlechtsspezifischen Unterschieden bezüglich der Personenmerkmale (Selbstkonzepte und Selbstwirksamkeitserwartung) im Fach Physik ausgegangen. Darüber hinaus könnte sich auch in der Experimentalgruppe das Geschlecht auf die erhobenen Variablen auswirken, da es Hinweise gibt, dass Mädchen und Frauen generell ein geringeres Interesse an Computerspielen zeigen (R. M. Brown et al., 1997; Cassell & Jenkins, 1998).

Auch die Schulform, Realschule oder Gymnasium, floss als Variable ein, da jeweils unterschiedliches Vorwissen vermutet wird (Realschule = 0, Gymnasium = 1).

Für die Bearbeitung der Materialien im Rahmen der Intervention waren gute Sprachkenntnisse erforderlich, da Informationen aus Texten entnommen und auch die Aufgabentexte verstanden werden mussten (Deutsch = 0, andere Erstsprache = 1).

4.1.9 Zusammenstellung der Instrumente

Alle Items wurden auf der Erhebungsseite *Sosci-Survey* (Leiner, 2023) eingepflegt. Insgesamt wurden zwei Pretests (A und B) sowie zwei Posttests (A und B) erstellt. Die A- und die B-Testformen unterschieden sich in den Fachwissenstest-Items. Außerdem wurden im Pretest, anders als im Posttest, das physikbezogene und das computerbezogene Selbstkonzept erhoben. Für die Erhebung der Daten des Pre- und des Posttests wurden Tablets eingesetzt, auf denen die Erhebungsseite *SoSci-Survey* per QR-Code aufgerufen werden konnte (Anhang C).

Durch eine einheitliche Methodik sowohl bei der Erhebung als auch bei der Auswertung, wie sie durch digitale Fragebogenformate gegeben ist, sollte die Testobjektivität sichergestellt werden (Brähler et al., 2002). Weiterhin soll die Durchführung der Tests auf Tablets ein einheitliches Vorgehen bei der Bearbeitung sowie zusätzliche Objektivität bei der Auswertung gewährleisten. Die Testobjektivität bezieht sich auf den Grad, in dem die Ergebnisse eines Tests frei von Verzerrungen sind und ausschließlich auf der Leistung des Testteilnehmers basieren (Moosbrugger & Kelava, 2020). Zunächst wurden die soziodemographischen Variablen, Geschlecht, Alter, Schulform sowie Erstsprache, abgefragt. Anschließend mussten die Teilnehmenden die Skalen der Selbstkonzepte sowie der Selbstwirksamkeitserwartung und zuletzt den Fachwissenstest ausfüllen. Sie wurden durch die Seite dazu aufgefordert, alles auszufüllen, sodass es keine fehlenden Werte gab. Die Fragebögen und die Interventionsmaterialien wurden zunächst pilotiert, um Erkenntnisse in Bezug auf die Länge der Intervention sowie die Güte der Fragebögen zu erlangen. Die Ergebnisse sind in Ergebnisse der Pilotierung 5.1 dokumentiert.

4.2 Statistische Methoden

In diesem Abschnitt werden die statistischen Verfahren, ihre Parameter und ihre Cutoff-Werte vorgestellt, mit denen die Hypothesen in dieser Arbeit untersucht wurden. Ziel ist es, Unterschiede zwischen verschiedenen Gruppen und Zeitpunkten zu ermitteln sowie Modelle zur Vorhersage zu überprüfen.

4.2.1 Ermittlung von Unterschieden zwischen zwei Gruppen oder Zeitpunkten

Zur Ermittlung der Unterschiede zwischen zwei Gruppen zum selben Zeitpunkt kann der t-Test bei unabhängigen Stichproben eingesetzt werden.¹⁷ Mit diesem Test können Unterschiede zwischen den Interventionsgruppen, aber auch zwischen beiden Geschlechtern ermittelt werden. Der t-Test für gepaarte Stichproben wird verwendet, um die Mittelwerte derselben Gruppe zu zwei Zeitpunkten zu erfassen, bspw. um die Pretest- und die Posttest-Ergebnisse zu vergleichen. Allgemein wird durch den t-Test ermittelt, ob es einen signifikanten Unterschied zwischen den Mittelwerten zweier Gruppen gibt. Der t-Test basiert auf der Annahme, dass die Daten normalverteilt sind. Berechnet wird die Differenz zwischen den Mittelwerten der beiden Gruppen und den Standardfehlern der Differenz.

¹⁷ Auch eine Varianzanalyse (ANOVA) ist eine Möglichkeit mit hoher statistischer Power, jedoch kommen bei zwei Gruppen ANOVA und t-Test zu gleichen Ergebnissen, wobei der Welch t-Test den Vorteil hat, einseitig und beidseitig testen zu können.

Schließlich wird diese Differenz mit einem t-Wert verglichen, der aus den Stichprobendaten berechnet wird, um die Wahrscheinlichkeit zu bestimmen, dass die Differenz zufällig ist. Sowohl beim gepaarten als auch beim ungepaarten Test wird zweiseitig auf Signifikanz getestet. Die Voraussetzung der Normalverteilung wurde mit dem Shapiro-Wilk-Test geprüft (Shapiro & Wilk, 1965). Da im vorliegenden Fall jedoch mit Stichprobengrößen-empfindlichem Verhalten zu rechnen war, wurde weiterhin mit dem D'Agostino-Test auf Schiefe der Verteilung geprüft. Dabei ist ein Test-Ergebnis innerhalb von $-.5$ und $.5$ ein Hinweis auf Symmetrie, wodurch die Annahme der Normalverteilung bestätigt werden kann (D'Agostino & Belanger, 1990). Weitere Auskunft gibt der Anscombe-Glynn-Test der Kurtosis (Anscombe & Glynn, 1983), wobei ein Ergebnis um 3 eine Normalverteilung repräsentiert. Eine weitere Voraussetzung ist die Annahme, dass die Varianzen der beiden Gruppen gleich sind (Homoskedastizität), was mit dem Levene-Test überprüft wird (O'Brien, 1978). Wenn dies nicht erfüllt ist, wird empfohlen, auf Welch's t-Test zurückzugreifen (Rasch et al., 2011). Manche Autor:innen raten dazu, Welch's t-Test grundsätzlich dem t-Test vorzuziehen (Ruxton, 2006). Es handelt sich um eine Variation des t-Tests, der robuster ist, wenn die Varianzen der beiden zu vergleichenden Gruppen signifikant unterschiedlich sind. In beiden Tests wird davon ausgegangen, dass die Daten normalverteilt und die beiden zu vergleichenden Gruppen unabhängig sind. Bei allen Tests wird ein kleiner p -Wert (üblicherweise kleiner als $.05$) als Zeichen für die Unwahrscheinlichkeit gedeutet, dass der Unterschied in den Medianen zwischen den beiden Gruppen auf Zufall zurückzuführen ist. Die Effektstärke wird bei den t-Tests mit Cohens d berechnet (Cohen, 2013). Hierfür wird die Differenz der Mittelwerte durch die Differenz der Standardabweichung geteilt. Werte von $.2$ bis $.4$ deuten auf schwache Effektstärken, Werte von $.5$ bis $.7$ auf mittlere und Werte $> .8$ auf starke Effektstärken hin (Fritz et al., 2012).

4.2.2 Vorhersage durch Regressionen

Die Regression wird verwendet, um die Beziehung zwischen einer abhängigen Variablen und einer oder mehreren unabhängigen Variablen zu untersuchen. Sie dient dazu, Hypothesen über die Beziehung zwischen Variablen zu testen und, basierend auf den Werten der unabhängigen Variablen, Vorhersagen über den Wert der abhängigen Variablen zu treffen. Mit Regressionsanalysen wurden die Hypothesen der zweiten Fragestellung (siehe Kapitel 3) überprüft. Die Regression wurde einerseits eingesetzt, um bestimmte (lineare) Beziehungen zwischen Variablen des Pretests und des Flow-Erlebens zu ermitteln. Andererseits diente sie auch der Erstellung des Strukturgleichungsmodells (siehe nächster Abschnitt), um die Einflussgrößen auf die Variablen des Posttests zu ermitteln. In diesem Abschnitt wird ein Überblick über die multiple lineare Regression sowie die moderierte und mediierte Regression, einschließlich der Annahmen des Modells, der Interpretation der Regressionskoeffizienten und der Bewertung der Modellanpassung, gegeben.

4.2.2.1 *Multiple lineare Regression*

Die multiple lineare Regression ist eine Form der linearen Regression, bei der mehrere unabhängige Variablen in das Modell aufgenommen werden können. Auf diese Weise wurden in der vorliegenden Arbeit die Einflüsse von Geschlecht, Selbstkonzepten, Selbstwirksamkeitserwartung(t_1) und

Fachwissen(t_1) auf das Flow-Erleben ermittelt.¹⁸ Die Grundgleichung für die multiple lineare Regression lautet: $Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + e$. Dabei ist Y die abhängige Variable, b_0 der Schnittpunkt, b_1 bis b_k sind die Regressionskoeffizienten, X_1 bis X_k die unabhängigen Variablen und e ist der Fehlerterm.

Im Folgenden werden die Voraussetzungen der multiplen linearen Regression aufgeführt.

- Linearität: Die Beziehung zwischen den unabhängigen und abhängigen Variablen der Regressionsgleichung ist linear. Dies kann grafisch durch paarweises Auftragen in einem Koordinatensystem sowie rechnerisch durch den Rainbow-Test (Utts, 1982) überprüft werden.
- Unabhängigkeit der Residuen: Die Beobachtungen im Datensatz sind voneinander unabhängig, was durch eine Zufallsstichprobe gewährleistet wird. Mithilfe des Durbin-Watson-Tests wird überprüft, ob die Residuen autokorreliert sind (Durbin & Watson, 1950).
- Normalverteilung der Residuen: Die Überprüfung kann einerseits visuell mit Histogrammen durchgeführt werden und mit dem Shapiro-Wilk-Test, bei dem Signifikanz gegen eine Normalverteilung spricht, überprüft werden (Shapiro & Wilk, 1965). Weiterhin kann die Schiefe, die nahe 0 sein muss, mit dem D'Agostino Test überprüft werden, bei dem Signifikanz gegen eine Normalverteilung spricht (D'Agostino & Belanger, 1990). Auch die Kurtosis nahe 3, die mit dem Anscombe Test geprüft werden kann, spricht für eine Normalverteilung, wobei auch hier der Test nicht Signifikant sein darf (Anscombe & Glynn, 1983).
- Homoskedastizität der Residuen: Die Varianz des Fehlerterms ist über alle Ebenen der unabhängigen Variablen hinweg konstant. Dies kann sowohl visuell (möglichst unstrukturierte Punktwolke) als auch mit dem Breusch-Pagan-Test (Signifikanz deutet auf Heteroskedastizität hin) überprüft werden (Breusch & Pagan, 1979).
- Abwesenheit von Multikollinearität: Die Prädiktoren korrelieren nicht miteinander. Da beim Selbstkonzept und der Selbstwirksamkeitserwartung möglicherweise auf der Ebene spezifischer Bereiche wie der physikbezogenen Einstellungen ähnliche Konstrukte abgefragt werden (Bong & Skaalvik, 2003), könnten auch in dieser Arbeit hohe Korrelationen zwischen diesen beiden Konstrukten auftreten. Auch die Selbstkonzepte als solche können möglicherweise korrelative Zusammenhänge aufweisen, da Schüler:innen mit einer positiven computerbezogenen Einstellung möglicherweise auch interessierter an MINT-Fächern wie Physik sind (Giammarco et al., 2015). Ein weiterer Grund ist, dass beide Konstrukte in der Hierarchie des akademischen Selbstkonzepts im mathematisch-naturwissenschaftlichen Bereich eingeordnet werden. Daher erhielt diese Prüfung besondere Aufmerksamkeit. Einerseits wurde durch eine zweiseitige Pearson-Korrelation der Korrelationskoeffizient ermittelt, der unterhalb der von Field et al. (2012) empfohlenen Grenze von $r = .8$ liegen sollte. Des Weiteren kann Multikollinearität am *variance influence factor (VIF)* abgelesen werden, wobei Werte über 10 kritisch bewertet werden (Kutner et al., 2004).
- Keine Ausreißer im Datensatz: Dies kann durch Prüfung der studentisierten Residuen kontrolliert werden, welche den Unterschied zwischen beobachteten und prognostizierten Zielwerten vergleichen (Abweichung sollte nicht > 3 sein) (Chambers et al., 2002). Weitere Sicherheit gibt die Cook-Distanz, die überprüft, ob und wie stark Ausreißer im Datensatz die

¹⁸ In manchen Arbeiten die hierarchische bzw. Schrittweise Regression zum Einsatz, bei der mittels eines iterativen Algorithmus die Kombination aus abhängigen Variablen ermittelt wird, die die unabhängige Variable am besten erklärt (Eid et al., 2013). Diese zeigt ihre Vorteile, wenn das Regressionsmodell nicht theoriegeleitet aufgestellt wird, sondern stellt ein exploratorisches Verfahren dar (ebd.). Da in dieser Arbeit das Regressionsmodell theoriegeleitet aufgestellt wird, kommt dieses Verfahren hier nicht zum Einsatz.

Regression beeinflussen (als Daumenregel gilt, dass einflussreiche Beobachten bei einer Cook-Distanz > 1 auftreten) kontrolliert werden (Cook & Weisberg, 1995).

Um die Regressionskoeffizienten zu interpretieren, werden üblicherweise die p -Werte untersucht, die jeder unabhängigen Variablen zugeordnet sind. Ein p -Wert misst die Wahrscheinlichkeit, dass der Koeffizient nicht von null abweicht. Ist er kleiner als ein Signifikanzniveau, wird der Einfluss der unabhängigen Variable als statistisch signifikant betrachtet.

Um die Güte des Regressionsmodells zu bewerten, werden mehrere Maße verwendet, z. B. das Bestimmtheitsmaß R^2 bzw. korrigiertes R^2 . Dabei ist R^2 ein Maß dafür, wie gut das Regressionsmodell an die Daten angepasst ist und wie viel der Streuung der Daten (Varianz) durch das Regressionsmodell erklärt werden. Dabei zeigen Werte nahe 1 eine gute Anpassung (100%) und Werte nahe 0 eine schlechte Anpassung an. Eine Kritik an R^2 ist seine Empfindlichkeit gegenüber der Anzahl der Variablen: Mit zunehmender Anzahl der Variablen wird auch R^2 größer. Das angepasste R^2 ist eine modifizierte Version von R^2 , bei der die Anzahl der unabhängigen Variablen im Modell berücksichtigt wird und entsprechende Korrekturen vorgenommen werden. Die Höhe von R^2 variiert je nach Disziplin: In exakten Fachgebieten wie den Naturwissenschaften gehen die Werte von R^2 in Richtung 1, in den Sozialwissenschaften sind Werte von .05 bis .50 üblich. In den Naturwissenschaften sind die Beziehungen zwischen Variablen oft gut definiert und konsistent, was zu höheren R^2 -Werten führt. In den Sozialwissenschaften sind die Beziehungen zwischen Variablen oft komplexer und weniger konsistent, was zu niedrigeren R^2 -Werten führt. Entsprechend muss die Interpretation immer in Abhängigkeit vom Kontext erfolgen (Schlittgen, 2013).

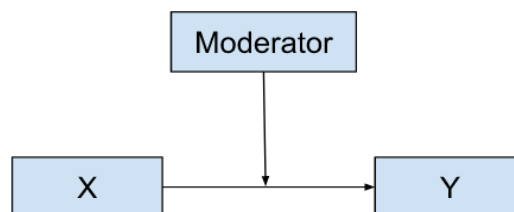
4.2.2.2 Moderierte Regression

Moderation ist eine statistische Technik, mit der untersucht wird, wie sich die Beziehung zwischen einer unabhängigen Variablen X und einer abhängigen Variablen Y in Abhängigkeit von einer Moderatorvariable M verändert. Das Pfaddiagramm ist in

Abbildung 16 zu sehen. In der vorliegenden Arbeit wurde auf diese Weise die Rolle des Flow-Erlebens im Strukturzusammenhang mit dem Fachwissen(t_1) und dem Fachwissen(t_2) bzw. der Selbstwirksamkeitserwartung(t_1) und der Selbstwirksamkeitserwartung(t_2) untersucht.

Bei einer Moderation ist die Wirkung der unabhängigen Variablen X auf die abhängige Variable Y nicht über alle Ebenen der Moderatorvariablen hinweg konstant. Die signifikanten Interaktionsbereiche können durch Johnson-Neyman-Intervalle gezeigt werden. Zur besseren Interpretation des möglichen Interaktionseffektes ist eine Zentrierung der Variablen zu empfehlen, auch wenn dies am Ergebnis nichts ändert (Kromrey & Foster-Johnson, 1998).

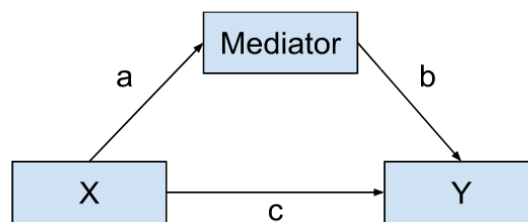
Abbildung 16 Pfaddiagramm der Moderation.



4.2.2.3 Mediierte Regression

Mediation ist eine statistische Technik zur Identifizierung des Mechanismus, wodurch eine unabhängige Variable eine abhängige Variable beeinflusst. Bei der Mediation wird eine dritte Variable, Mediator genannt, in das Modell aufgenommen und verwendet, um die Beziehung zwischen der unabhängigen und der abhängigen Variablen zu erklären. Die Mediatorvariable liegt zwischen der unabhängigen Variablen und der abhängigen Variablen und wird als kausale Verbindung zwischen ihnen betrachtet (siehe Abbildung 17). Mit der mediierten Regression wurde als alternative Theorie die Rolle des Flow-Erlebens bei der möglichen Veränderung von Fachwissen und Selbstwirksamkeitserwartung untersucht.

Abbildung 17 Pfaddiagramm der Mediation.



4.2.3 Untersuchungen mit Strukturgleichungsmodellierung

Mit der Strukturgleichungsmodellierung, auf Englisch *Structural Equation Modeling* (SEM), werden die Beziehungen zwischen mehreren Variablen analysiert, um komplexe Hypothesen zu prüfen (Werner, 2015). Mit der SEM können Modelle entwickelt werden, die dazu dienen, die kausalen Beziehungen zwischen beobachteten und latenten Variablen aufzuzeigen sowie die Stärke und Signifikanz der Beziehungen zu testen. Die SEM ermöglicht es auch, die Gültigkeit eines theoretischen Modells zu testen, indem seine Anpassung an die Daten untersucht wird (ebd.). Zur Beantwortung der dritten Forschungsfrage wurden die Variablen mithilfe der SEM in Zusammenhang gebracht und untersucht. Die Grundlage der SEM ist die Regressionsanalyse, um die Auswirkungen einer oder mehrerer Prädiktorvariablen auf eine Reihe von Ergebnisvariablen zu schätzen. Im Gegensatz zur traditionellen Regressionsanalyse ermöglicht die SEM jedoch die Schätzung sowohl direkter als auch indirekter Effekte zwischen Variablen. Eine der Hauptstärken der SEM ist ihre Fähigkeit, latente Variablen zu modellieren. Diese werden nicht direkt beobachtet, sondern stattdessen aus einer Reihe von beobachteten (manifesten) Variablen abgeleitet, zusammen ergibt dies das Messmodell. Im Messmodell kann der einzelne Beitrag der manifesten Variablen zur latenten Variable dargestellt werden und so ermittelt werden, welchen Anteil der Varianz sie erklären können.

Für die SEM müssen nach Urban und Mayerl (2018) neben den Bedingungen der Regressionsanalyse Unabhängigkeit und Linearität bestimmte Voraussetzungen erfüllt sein, die im Folgenden aufgeführt werden.

- Identifizierbarkeit: Das Modell ist identifizierbar, d. h., die Parameter können eindeutig aus den beobachteten Daten geschätzt werden (ebd.).
- Positive Bestimmtheit: Die Kovarianzmatrix der beobachteten Variablen ist positiv bestimmt (ebd.).
- Stichprobengröße: Die Stichprobe muss groß sein, um sicherzustellen, dass das Modell identifizierbar ist und die Parameter genau geschätzt werden können. Vorsicht ist geboten, wenn die Stichprobe klein ist oder die Daten nicht normalverteilt sind (ebd.).

Die Passung des geschätzten Modells mit den Daten wird durch verschiedene Modellparameter bestimmt (Hu & Bentler, 1999). Beispielsweise wird mit dem χ^2 -Test die Nullhypothese überprüft, ob

die Kovarianzmatrix des Modells mit jener der gemessenen Daten zusammenpasst, d. h., sein Ergebnis darf nicht signifikant sein mit Werten von $p > .05$ (Hu & Bentler, 1999, S. 2). Der Grund dafür ist, dass der χ^2 -Test unter anderem von der Stichprobengröße abhängt und mit großer Stichprobe häufig signifikant ausfällt (ebd.). Alternativ kann das Verhältnis von χ^2 zu den Freiheitsgraden, auf Englisch *Degrees of Freedom* (*df*) herangezogen werden, bei dem Werte ≤ 2 als gut und solche ≤ 3 als akzeptabel betrachtet werden (Bollen, 1989, S. 278). Ein weiteres Gütekriterium stellt der mittlere quadratische Fehler, auf Englisch *Root Mean Square Error of Approximation* (RMSEA) dar. Er misst die Diskrepanz zwischen dem Modell und den gemessenen Werten, wobei Werte $< .05$ (G. T. Brown, 2006) bzw. $< .08$ (Browne & Cudeck, 1993) eine bessere Anpassung repräsentieren. Der RMSEA reagiert empfindlich bei kleinen DF und kleinem Stichprobenumfang, sodass auch korrekt spezifizierte Modelle verworfen werden können (G. T. Brown, 2006). Die standardisierte Quadratwurzel des mittleren Residuenquadrats, auf Englisch *Standardized Root Mean Square Residual* (SRMR), ist ein Index für den Durchschnitt der standardisierten Residuen zwischen den beobachteten und den hypothetischen Kovarianzmatrizen (F. F. Chen, 2007). Dabei gelten Werte $< .1$ als akzeptabel (Hu & Bentler, 1999). Auch SRMR ist nachweislich empfindlich gegenüber dem Stichprobenumfang (Hu & Bentler, 1999). Die letzten beiden Gütekriterien sind der *Comparative Fit Index* (CFI) und der *Tucker Lewis Index* (TLI). Beide vergleichen das Modell mit einem Basismodell, wobei Werte $\geq .95$, Werte möglichst nahe 1, aber auch Werte im Bereich von .90 bis .95 auf eine akzeptable Modellanpassung hinweisen können (Bentler, 1990).

Tabelle 12 Cutoff-Werte für Strukturgleichungsmodelle.

Name	Autor/Werk	Akzeptabler Fit	Guter Fit
Verhältnis χ^2/df	Bollen, 1989	$\chi^2/df \leq 3$	$\chi^2/df \leq 2$
RMSEA	Browne & Cudeck, 1993	$\leq .08$	$\leq .05$
SRMR	Hu & Bentler, 1999	$\leq .10$	$\leq .05$
CFI und TLI	(Bentler, 1990)	$\geq .90$	$\geq .95$

Anmerkungen. SRMR: Standardized Root Mean Square Residual, RMSEA: Root Mean Square Error of Approximation, CFI: Comparative Fit Index, TLI: Tucker Lewis Index.

Die Angaben in Tabelle 12 sind als grobe Orientierung aufzufassen (Hu & Bentler, 1999), da die empfohlenen Cutoff-Werte immer vom Modell abhängen. Bei der Betrachtung der Ergebnisse ist jedoch Vorsicht geboten, denn auch Daten, die mit dem Modell gut zusammenpassen, implizieren nicht zwingend, dass dieses der Realität entspricht (Bollen 1989). Zusätzliche Robustheit ermöglicht Bootstrapping sowie die Bollen-Stine-Bootstrapping-Methode (Bollen & Stine, 1992), bei der das Modell auf Stichprobendaten übertragen wird. Zur Beurteilung der Modelle können unterschiedliche Algorithmen verwendet werden. Häufig kommt dabei ein robustes Maximum-Likelihood-Schätzverfahren zum Einsatz (Satorra & Bentler, 1994), das als zuverlässig gilt (Tabachnick & Fidell, 2007). Zu den Voraussetzungen für die Maximum Likelihood Methode gehört, dass die exogenen Variablen sowie die Fehlerterme unabhängig sind, die endogenen Variablen multivariat normalverteilt sind und das Modell korrekt spezifiziert wurde. Ein Modell, bei dem mehr Informationen bekannt als unbekannt sind, gilt als identifiziertes Modell: Alle Parameter des Modells können eindeutig geschätzt werden. Die Ausprägung der Identifizierung (unteridentifiziert, gerade identifiziert und überidentifiziert) kann über die Freiheitsgrade bestimmt werden. Ein Modell ist identifiziert, wenn es genau eine Lösung für das Gleichungssystem gibt, während es nicht identifiziert ist, wenn mehrere gleichwertige Lösungen existieren (Backhaus et al., 2021).

Eine Möglichkeit, insbesondere bei kleiner Stichprobe, die Anzahl der zu schätzenden Parameter zu reduzieren, stellt das Item-Parceling dar, bei dem Items zu Päckchen zusammengefasst werden, was zu weniger Fehlervarianzen, Varianzen, Kovarianzen und Faktorladungen führt (Bandalos & Finney, 2002), was jedoch auch zu einem Verlust von Information führen kann (ebd.). Die Päckchen können beispielsweise anhand der Faktorladungen der Items einer Skala zusammengefasst werden (ebd.).

4.2.4 Verwendete Software und Bibliotheken

Für die Auswertung wurde ausschließlich mit der Statistiksoftware R gearbeitet (RStudio Team, 2023). Die Reliabilitätsprüfung anhand von McDonalds ω wurde mit der Bibliothek *psych* (Revelle & Zinbarg, 2009) durchgeführt. Für den t-Test und Welch's t-Test sowie Cohens d wurden die Methoden der Bibliothek *misty* (Rasch et al., 2011) verwendet. Die Voraussetzungen wurden mit Methoden der Bibliotheken *moments* (Komsta & Novomestky, 2015) bzw. *lawstat* (Gastwirth et al., 2017) geprüft. Für moderierte und medierte Regressionsmodelle wurden die Methoden der *PROCESS*-Bibliothek von (Hayes, 2017) verwendet, die auf dem Prinzip der kleinsten Quadrate (Draper & Smith, 1998) beruhen, um nicht standardisierte Koeffizienten und heteroskedastizitätskonsistente Standardfehler (HC3, Davidson & MacKinnon, 1993) zu ermitteln. Zusätzliche Robustheit wurde durch Bootstrapping mit 5000 Iterationen erreicht. Die Prüfung der Voraussetzungen der Regression erfolgte mit den Methoden der *olsrr*-Bibliothek. Für die multiplen linearen Regressionen sowie Strukturgleichungsmodelle wurde mit den Methoden aus *lavaan* gearbeitet. Zusätzliche Robustheit wurde durch die Bollen-Stine-Methode aus *lavaan* (Rosseel, 2012) erreicht.

4.2.5 Statistische Power und Stichprobengröße

Es wurden 8. Klassen ohne Vorwissen in der Elektrizitätslehre an Gymnasien und Realschulen in Baden-Württemberg rekrutiert. Schüler:innen der Sekundarstufe scheinen besonders von Serious Games zu profitieren und erzielen höhere Lernerfolge als jüngere Schüler:innen (Vogel et al., 2006; Wouters et al., 2013). Des Weiteren wird vermutet, dass in diesem Alter das Interesse an MINT-Fächern, insbesondere bei Mädchen, abnimmt (Solga & Pfahl, 2009).

Eine A-priori-Power-Analyse wurde mit *G*Power*, Version 3.1.9.7, (Faul et al., 2009) durchgeführt, um die minimale Stichprobengröße zu bestimmen, die zum Testen der Studienhypothesen erforderlich ist. Die Ergebnisse zeigten, dass die erforderliche Stichprobengröße zum Erreichen einer Trennschärfe von 80 % zum Erkennen eines mittleren Effekts bei einem Signifikanzkriterium von $p = .05$ eine Gruppengröße von $N = 107$ für die multiple Regressionsanalyse und $N = 100$ für die anteiligen Gruppen des t-Tests ist.

Hinsichtlich eines angemessenen Stichprobenumfangs für die SEM besteht in der Literatur kein Konsens, jedoch gibt es einige Hinweise, dass einfache SEM-Modelle auch bei kleinen Stichproben sinnvolle Ergebnisse liefern (Hoyle, 1999). Teilweise fordern Wissenschaftler eine Stichprobe von mindestens $N = 150$ Beobachtungen, wobei diese Zahl durch Simulationsstudien bestätigt werden konnte (Muthén & Muthén, 2002). Andere Forschende empfehlen für die Modellierung als Faustregel $N = 100$ Beobachtungen pro Gruppe bzw. mindestens $N = 200$ Beobachtungen (Kline, 2005). Die Stichprobengröße wird häufig im Hinblick auf die Anzahl der beobachteten Variablen betrachtet. Für normalverteilte Daten schlagen Bentler und Chou (1987) vor, dass fünf Beobachtungen pro Variable ausreichend ist, wenn latente Variablen mehrere Indikatoren haben. Eine weithin akzeptierte Regel sind zehn Fälle bzw. Beobachtungen pro Indikatorvariable, um eine untere Grenze für eine angemessene Stichprobengröße festzulegen (Charter, 1999). Zusammenfassend gilt für die SEM in dieser Arbeit, dass so viele Beobachtungen wie möglich, mindestens jedoch $N = 150-200$ gefordert werden.

4.3 Studiendesign und Durchführung

Das Ziel der vorliegenden Arbeit ist es, die Wirkungen eines Serious Games auf das Fachwissen sowie die Selbstwirksamkeitserwartung zu untersuchen. Die vermutete vermittelnde Funktion des Flow-Erlebens und der Zusammenhang mit den Selbstkonzepten sollen theoriegeleitet erklärt werden. Die vorliegende Untersuchung wurde als Interventionsstudie mit Kontrollgruppe durchgeführt (siehe Abschnitt 2.1.2), bei der die Kontrollgruppe ein Lernsetting¹⁹ erhielt, das die gleichen Inhalte wie das Serious Game umfasste, inklusive Aufgaben und Versuche. Durch die Kontrolle möglichst vieler Variablen soll Aufschluss darüber gewonnen werden, wie Schüler:innen, über das Fachwissen hinaus, auch bezüglich ihrer Selbstwirksamkeitserwartung gefördert werden können. Dieses Design ermöglicht darüber hinaus eine Vergleichbarkeit der Bedingungen und Variablenkontrolle und bietet eine hohe Validität und die Möglichkeit kausaler Schlussfolgerungen.

Aus allen Items wurden auf *Sosci-Survey* (Leiner, 2023) zwei Fragebögen erstellt, die sich in den Fachwissensitems unterscheiden. Das Untersuchungsdesign ist in der Tabelle 13 dargestellt. Die Intervention dauerte drei Stunden, davon wurden ca. 15 Minuten vor und nach der Intervention für die Beantwortung der Fragebögen benötigt. Zu Beginn wurde die Klasse begrüßt und das Vorgehen kurz erklärt. Die Schüler:innen zogen daraufhin ein Los mit ihrem QR-Code, der mit den iPads gescannt wurde. Die Schüler:innen wurden gebeten, ihren Code über die Dauer der drei Stunden gut aufzubewahren, was auch, bis auf zwei Ausnahmen, in der gesamten Intervention gut funktionierte. Mittels des QR-Codes bekamen die Schüler:innen zufällig einen der Fragebögen, A oder B, zugewiesen, wurden einer der beiden Interventionsgruppen zugeordnet und erhielten eine fünfstellige Zufallszahl, durch die Materialien und Fragebögen eindeutig zugeordnet werden konnten. Nachdem die Fragebögen ausgefüllt worden waren, wurden die Interventionsgruppen getrennt. Die Schüler:innen der Experimentalgruppe erhielten in einem Computerraum USB-Sticks mit dem Serious Game als portable Version, die Proband:innen der Kontrollgruppe bekamen das entsprechende Material im Physikraum. Während der Intervention wurde das Flow-Erleben erhoben. In der Experimentalgruppe geschah dies im Spiel über Pop-up-Fenster, in der Kontrollgruppe mithilfe eines Fragebogens, der ausgefüllt wurde, wenn die Testleitung zu randomisierten Zeitpunkten dazu aufforderte.

Tabelle 13 Untersuchungsdesign.

Gruppe:	Experimentalgruppe		Kontrollgruppe	
Dauer:				
15 min	Fragebogen A	Fragebogen B	Fragebogen A	Fragebogen B
105 min	Computerraum: Serious Game	Flow-Kurz-Skala	Physikraum: Materialkiste	Flow-Kurz-Skala
15 min	Fragebogen B	Fragebogen A	Fragebogen B	Fragebogen A

Am Ende der Intervention füllten beide Gruppen die zweiten Fragebögen aus. Die Gruppe A erhielt die Fragebögen B und *vice versa*. Zudem wurden die USB-Sticks sowie die Materialkisten der Kontrollgruppe eingesammelt. Die Daten von *Sosci-Survey*, aus dem Spiel und den Papierbögen der Kontrollgruppe wurden aufbereitet, codiert und anschließend analysiert. Im Folgenden werden die Ergebnisse der Pilotierung sowie der Erhebung erläutert und anschließend diskutiert.

¹⁹ Aufgabenblätter, Infotexte und ein einfaches Stecksystem für Versuche zur Elektrizitätslehre, wie z. B. Multimeter, Glühlampen und Batterien.

5 Ergebnisse

In diesem Kapitel wird die Überprüfung der aus der theoretischen Grundlage abgeleiteten Hypothesen anhand der erhobenen Daten und verschiedener Methoden der statistischen Datenanalyse vorgestellt. Es werden sowohl Methoden zur Prüfung von Unterschieden als auch Regressionsanalysen und Strukturgleichungsmodelle verwendet, um die theoretischen Modelle mit den Daten zu vergleichen. Zuerst werden die Erfahrungen und Ergebnisse der Pilotstudie präsentiert, gefolgt von den Ergebnissen der Haupterhebung.

5.1 Ergebnisse der Pilotierung

An der Pilotierung nahmen 142 Schüler:innen (69 männlich, 73 weiblich) aus 8. Klassen von Realschulen ($n = 69$) und Gymnasien ($n = 73$) in Baden-Württemberg teil. Die Wahl fiel auf diese beiden Schulformen, da sie im Bildungsplan übereinstimmen und durch ähnliche Unterrichtsvermittlungen eine höhere Variablenkontrolle ermöglichten. Die Schüler:innen waren durchschnittlich 13.35 Jahre alt ($SD = .56$). Weiter gaben 31 Schüler:innen an, zuhause kein Deutsch zu sprechen. Die Daten der Experimentalgruppe (79 Datensätze, 31 männlich, 48 weiblich) stammten von 38 Gymnasial- und 41 Realschüler:innen, wobei acht Schüler:innen des Gymnasiums und acht der Realschule aussagten, zuhause kein Deutsch zu sprechen. In der Kontrollgruppe (63 vollständige Datensätze, davon 38 männlich, 25 weiblich, 15 sprechen kein Deutsch zuhause) besuchten 35 Schüler:innen das Gymnasium, von denen acht angaben, zuhause kein Deutsch zu sprechen. Die Fragebögen zu Selbstkonzepten, Selbstwirksamkeitserwartung und Flow-Erleben sowie der Fachwissenstest wurden eingesetzt, dabei wurde Fragebogen A von 76 Schüler:innen bearbeitet. Die Interventionsdauer über drei Schulstunden zeigte sich als ausreichend, wobei einige Kinder in beiden Gruppen noch das Bonuskapitel bearbeiteten. Alle Ergebnisse sind in der Tabelle 14 zu sehen.

Tabelle 14 Ergebnisse der Pilotierung.

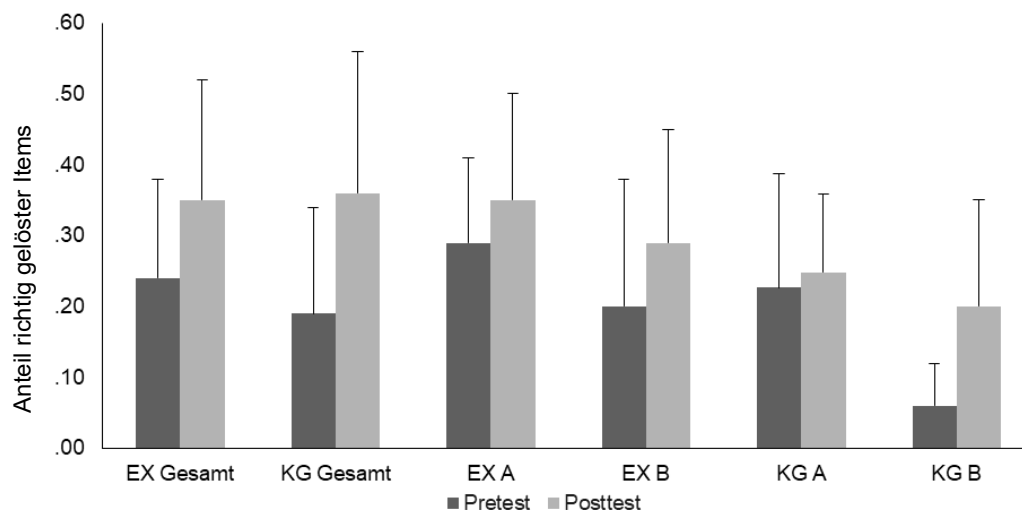
	Experimentalgruppe				Kontrollgruppe		
	ω	N	M	SD	N	M	SD
Gesamt	-	79			63		
Gruppe A	-	35			41		
Männlich	-	31			38		
Gymnasium	-	38			35		
Erstsprache D	-	16			15		
Alter	-		13.3	.46		13.4	.66
SKP	.86		3.08	.79		3.22	.85
SKC	.90		3.12	.89		3.58	.91
Flow	.81		3.21	.59		3.24	.83
FW _V	-		.24	.14		.19	.15
FW _N	-		.35	.17		.26	.20
SWE _V			2.87	.85		2.93	.88
SWE _N	.83		3.15	1.05		3.27	.94

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept; FW_V = Fachwissen(t_1), FW_N = Fachwissen(t_2), SWE_V = Selbstwirksamkeitserwartung(t_1), SWE_N = Selbstwirksamkeitserwartung(t_2).
 ω = McDonalds Omega, M = Mittelwert, SD = Standardabweichung.

5.1.1 Pilotierung des Fachwissenstests

Abbildung 18 zeigt die Ergebnisse der Fachwissenstests. Dabei ist der Vergleich der Interventionsgruppen miteinander sowie innerhalb der Interventionsgruppen zwischen den beiden Varianten A und B zu sehen. In beiden Gruppen sind signifikante Verbesserungen des Fachwissens erkennbar, sowohl in der Variante A als auch in der Variante B. In der Experimentalgruppe wurden im Fachwissenstest(t_1) vor der Intervention im Mittel $M = .24$ Punkte erzielt ($SD = .14$) und im Fachwissenstest(t_2) nach der Intervention $M = .35$ Punkte ($SD = .17$). Dies stellt einen signifikanten ($t(75) = -4.26, p < .001, 95\%-CI [-.16, -.06]$) Lernzuwachs mit schwacher Effektstärke ($d = .49$) dar. In der Kontrollgruppe wurden im Fachwissenstest(t_1) $M = .19$ Punkte ($SD = .15$) erzielt und im Fachwissenstest(t_2) $M = .26$ ($SD = .20$), was ebenfalls einem signifikanten Lernzuwachs ($t(61) = -3.33, p = .001, d = .42, 95\%-CI [-.11, -.03]$) entspricht. Der Unterschied in der Lernwirksamkeit zwischen den Gruppen war nicht signifikant ($t(133.88) = 1.27, p = .208$).

Abbildung 18 Ergebnisse im Fachwissenstest vor und nach der Pilotierung.



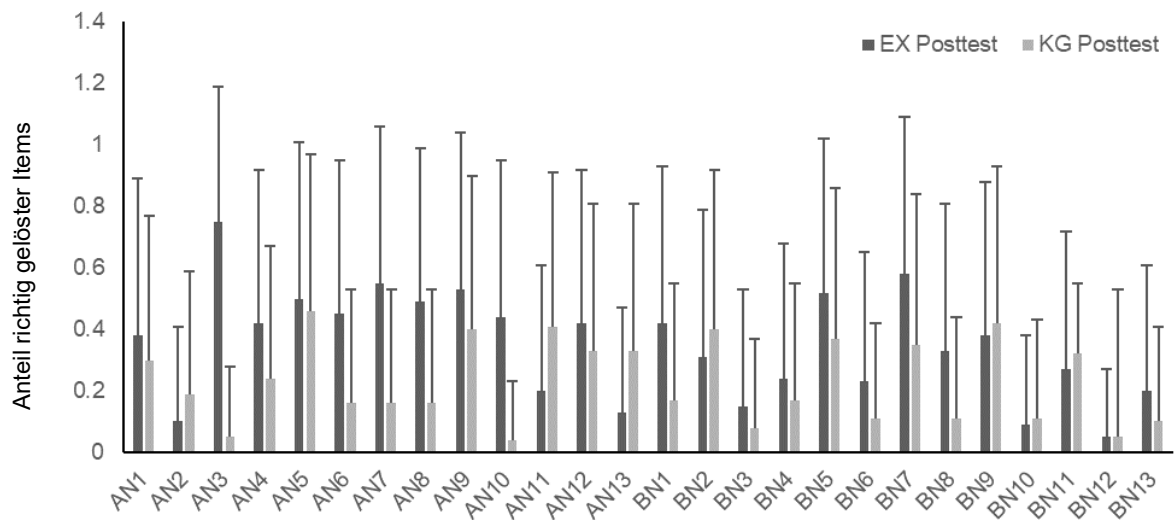
Anmerkungen. EX = Experimentalgruppe, KG = Kontrollgruppe.

In Abbildung 19 sind die Ergebnisse zu den einzelnen Items des A- und B-Tests dargestellt, weitere Ergebnisse sind in Anhang E. Die sogenannte Itemschwierigkeit stellt den prozentualen Anteil richtiger Lösungen bei einem bestimmten Item dar. Üblicherweise werden Items mit einer Lösungswahrscheinlichkeit unter 20 % und über 80 % ausgeschlossen (Mummendey & Grau, 2014). Um Boden- und Deckeneffekte die fehlende Berücksichtigung von Schüler:innen mit sehr viel oder sehr wenig Fachwissen zu verhindern, sollten die Items des Fachwissensbereichs ein möglichst großes Spektrum abdecken. Da es sich um ein etabliertes Testinstrument handelt und die Anzahl an Items gering war, wurde entschieden, auch Items unter 10 % zu berücksichtigen. Aufgrund der hohen Schwierigkeit wird der Fachwissenstest im Anschluss kritisch diskutiert.

Auffällig sind die insgesamt geringe Punktzahl bei beiden Tests sowie die geringere Punktzahl in der Variante B des Fachwissenstests. Dies kann einen Hinweis für eine allgemein hohe Schwierigkeit und eine höhere Schwierigkeit der Testvariante B darstellen. Jedoch sind zwei Dinge zu berücksichtigen: Zum einen war die Verteilung der Testhefte A und B auf die Gruppen nicht einheitlich. Weiterhin lag die Stichprobengröße weit unterhalb des geforderten Minimums von $N = 100$ pro Gruppe. Daher muss

die Schwierigkeit in der Hauptuntersuchung erneut geprüft und gegebenenfalls durch *Test-Equating* (Kolen & Brennan, 2014) angeglichen werden.²⁰

Abbildung 19 Ergebnisse des Fachwissenstests beider Interventionsgruppen für den A- und B-Test.



Anmerkungen. AN = Items der Gruppe A und BN = Items der Gruppe B im Fachwissenstest(t₂).

5.1.2 Pilotierung der Selbstwirksamkeitserwartung

Beide Gruppen konnten ihre Selbstwirksamkeitserwartung durch die Intervention signifikant verbessern (Experimentalgruppe: $t(77) = -3.87$, $p = .005$, $d = .33$; Kontrollgruppe: $t(62) = -4.15$, $p = .001$, $d = .49$) und unterscheiden sich in der Veränderung nicht voneinander ($t(138.67) = -.51$, $p = .613$). Die Reliabilitätsprüfung mit *McDonalds Omega* für die Skala lieferte mit $\omega = .83$ ein zufriedenstellendes Ergebnis. Die Ergebnisse für die einzelnen Items der Skalen finden sich in Anhang F.

5.1.3 Pilotierung des Flow-Erlebens

Die Reliabilitätsprüfung mit *McDonalds Omega* ist mit $\omega = .81$ zufriedenstellend. Bei den Ergebnissen zeigten sich keine Unterschiede zwischen den beiden Interventionsgruppen bezüglich des erlebten Flows ($t(108.03) = -.21$, $p = .832$). Ein Grund dafür könnte in einem Fehler des Serious Games liegen, wodurch sich Spielstände in den Folgestunden nicht mehr aufrufen ließen. Dies sorgte für viel Frustration seitens der Schüler:innen, die das Spiel daraufhin erneut starten mussten. Dieser Fehler wurde für die Haupterhebung korrigiert. Die Ergebnisse zu den einzelnen Items befinden sich im Anhang H.

5.1.4 Pilotierung des computerbezogenen und physikbezogenen Selbstkonzepts

Vor der Auswertung wurden die Items SKP2, SKP3 und SKP5 umcodiert, bei diesen Items zeigte sich eine hohe Punktzahl in einer negativen Ausprägung des Selbstkonzepts. In Bezug auf das physikbezogene Selbstkonzept gibt es keine Unterschiede zwischen den Gruppen, jedoch zeigte sich

²⁰ Im Falle ungleich schwerer Testvarianten des Fachwissenstests kann durch linear Equating und Linking ein vergleichbarer Mittelwert und eine vergleichbare Standardabweichung ermittelt werden. Beim linear Equating und Linking wird ein Rohwert auf die Testvariante 1 geschätzt, der jedem Rohwert der Testvariante 2 entspricht und so eine gemeinsame Messskala für beide Testvarianten erstellt, die die Schwierigkeitsunterschiede berücksichtigt (Kolen und Brennan (2014).

ein geringeres computerbezogenes Selbstkonzept in der Experimentalgruppe. Da diese Daten aus der Testung vor der Aufteilung in die einzelnen Interventionsgruppen stammen, ist hier von einem Zufall auszugehen. Die Reliabilitätsprüfung mit McDonalds Omega ergab für die Skala des physikbezogenen Selbstkonzepts mit $\omega = .86$ und für jene des computerbezogenen Selbstkonzepts mit $\omega = .90$ zufriedenstellende Ergebnisse. Die Resultate der einzelnen Items sind im Anhang G dargestellt.

5.2 Ergebnisse der Hauptstudie

Die Erhebung wurde im Schuljahr 2022/2023 durchgeführt. Insgesamt erklärten sich sechs Gymnasien und zwei Realschulen bereit, an der Erhebung teilzunehmen.²¹ Ein Datensatz gilt als vollständig, wenn Schüler:innen über die gesamte Dauer von drei Schulstunden anwesend waren und Pretest sowie Posttest ausgefüllt haben. An der Erhebung nahmen insgesamt 363 Schüler:innen teil. Da die Schüler:innen auf der Erhebungsseite *Sosci-Survey* zur vollständigen Bearbeitung der Fragebögen angehalten wurden, gab es keine fehlenden Einzelwerte, jedoch fehlende Pre- oder Post-Datensätze, bedingt durch hohe Abwesenheitszahlen während der COVID-Pandemie. Nachdem fehlende Datensätze listenweise ausgeschlossen wurden, verblieben insgesamt 310 vollständige Datensätze (158 bzw. 51 % männlich), mit einem Anteil von 236 Gymnasiast:innen (76 %). Der Altersdurchschnitt lag bei 13.32 Jahren ($SD = .54$).

Insgesamt wurden in der Hauptuntersuchung 168 vollständige Datensätze der Experimentalgruppe und 142 vollständige Datensätze der Kontrollgruppe berücksichtigt. Die deskriptiven Ergebnisse der Erhebung sind in Tabelle 15 zu sehen.

Bei erster Betrachtung werden Unterschiede unter anderem hinsichtlich des Fachwissens deutlich: Die Werte beim Fachwissen(t_1), aber auch beim Fachwissen(t_2) sind in der Kontrollgruppe geringer als in der Experimentalgruppe. Auch bei der Selbstwirksamkeitserwartung zeigten sich Unterschiede: Die Werte in der Experimentalgruppe sind sowohl für die Selbstwirksamkeitserwartung(t_1) als auch für die Selbstwirksamkeitserwartung(t_2) höher als in der Kontrollgruppe. Weiterhin fallen Unterschiede beim Flow-Erleben auf. Hier wies die Experimentalgruppe im Durchschnitt höhere Werte auf als die Kontrollgruppe. Ob die Unterschiede signifikant sind, soll in Abschnitt 5.2.5 untersucht werden. Zunächst sollen die Ergebnisse der Fragebögen dargestellt werden.

²¹ Durch die Pandemie wurde die Suche erschwert, lange Zeit fand der Unterricht ausschließlich im Homeschooling statt, dann gab es Wechselklassenmodelle, bei denen nur die halben Klassen in der Schule Unterricht erhielten, während die andere Hälfte per Video zugeschaltet wurde. Und selbst im Regelunterricht öffneten nur wenige Schulen ihre Türen für außenstehende Personen, um das Infektionsrisiko gering zu halten. Eine weitere Schwierigkeit stellte die nach wie vor mangelhafte Ausstattung mit Computern an Schulen dar.

Tabelle 15 Ergebnisse der Erhebung.

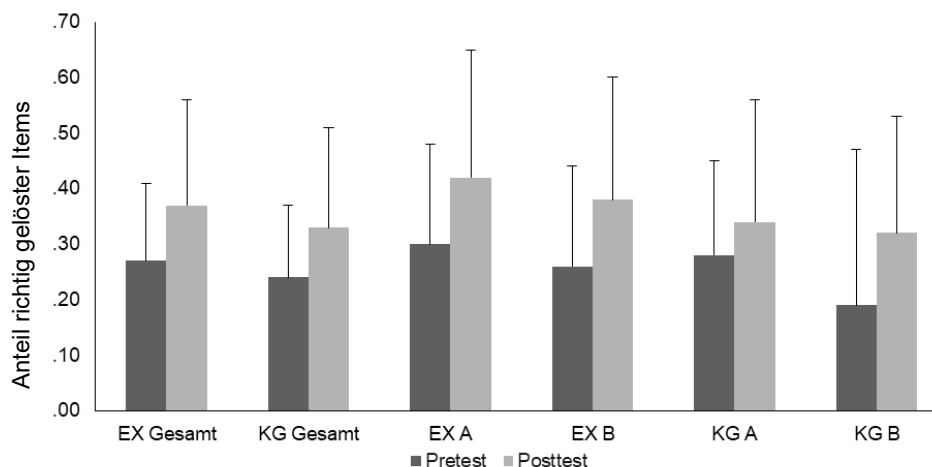
	Experimentalgruppe				Kontrollgruppe		
	ω	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
N	-	168			142		
Gruppe A	-	86			58		
Männlich	-	85			73		
Gymnasium	-	123			113		
Erstsprache D	-	144			110		
Alter	-		13.27	.52		13.39	.58
SKP	.80		3.06	.78		2.98	.79
SKC	.90		3.30	.93		3.36	.85
Flow	.82		3.43	.81		3.15	.71
FW _V	-		.27	.14		.24	.13
FW _N			.35	.19		.33	.18
SWE _V	.84		3.08	.86		3.06	.86
SWE _N			3.26	.80		3.16	.82

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept; FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂). ω = McDonalds Omega, *M* = Mittelwert, *SD* = Standardabweichung.

5.2.1 Fachwissen

Das Ergebnis der Pilotierung bezüglich unterschiedlicher Schwierigkeiten des A- und B-Tests konnten nicht reproduziert werden ($t(113) = 1.87, p = .064$), sodass eine Angleichung der Schwierigkeit nicht nötig war. Die vollständigen Ergebnisse befinden sich in Anhang I.

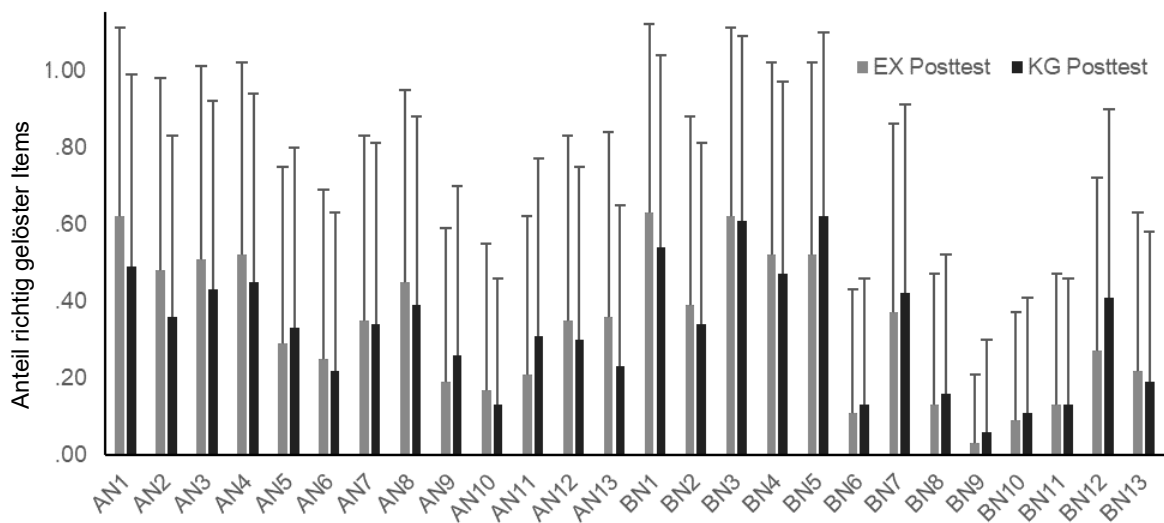
Abbildung 20 Gegenüberstellung der Pretest und Posttest Ergebnisse.



Anmerkungen. EX = Experimentalgruppe, KG = Kontrollgruppe.

Bei Betrachtung der Ergebnisse der Posttests in Abbildung 21 fällt jedoch auf, dass die höchsten durchschnittlichen Lösungswahrscheinlichkeiten bei unter 70% liegen, was auf eine hohe Schwierigkeit der Fragebögen schließen lässt. Bei mehreren Items (A10, B8 und B9) liegt die Lösungswahrscheinlichkeit sogar unter 10 %. Der Fragebogen scheint damit in seiner Schwierigkeit unangemessen für eine 8. Klasse, was in der Diskussion berücksichtigt wird.

Abbildung 21 Gegenüberstellung der Gruppen bei den Fachwissensergebnissen.



Anmerkungen. AN = Items der Gruppe A und BN = Items der Gruppe B im Fachwissenstest(t_2).

5.2.2 Selbstwirksamkeitserwartung

Bei Betrachtung der Items (siehe Tabelle 16) fällt die große Differenz beim Antwortverhalten zwischen der hohen Zustimmung für Item SV1 (*Ich kann auch die schwierigen Aufgaben in Physik lösen, wenn ich mich anstrengte*) und die geringere Zustimmung bei Item SV2 (*Wenn ich eine schwierige Physikaufgabe an der Tafel lösen soll, glaube ich, dass ich das schaffen werde*) in beiden Gruppen auf. Auch im Nachtest ist diese Differenz noch zu beobachten. Bei Betrachtung der Mittelwerte von SN5 fällt auf, dass hier insbesondere die Experimentalgruppe eine Verbesserung erzielt hat (Ich bin mir sicher, dass ich auch dann noch meine gewünschten Leistungen in Physik erreichen kann, wenn ich mal eine schlechte Note bekommen habe), während der Mittelwert in der Kontrollgruppe kleiner scheint. Eine genauere Überprüfung dieser Unterschiede soll im Abschnitt 5.2.5 stattfinden.

Tabelle 16 Ergebnisse zur Selbstwirksamkeitserwartung

Item	Pretest Experimentalgruppe		Pretest Kontrollgruppe	
	M	SD	M	SD
SV1	3.37	1.04	3.37	1.07
SV2	2.76	1.12	2.66	1.12
SV3	2.94	1.14	2.99	1.14
SV4	3.32	1.07	3.31	1.09
SV5	3.00	1.21	2.99	1.20
Gesamt	3.08	0.86	3.06	0.86
Item	Posttest Experimentalgruppe		Posttest Kontrollgruppe	
	M	SD	M	SD
SN1	3.57	1.00	3.52	1.01
SN2	2.96	1.00	2.94	1.05
SN3	3.07	1.04	3.09	1.03
SN4	3.36	1.06	3.31	1.05
SN5	3.08	1.14	2.96	1.19
Gesamt	3.25	0.80	3.16	0.82

Anmerkungen. SV = Selbstwirksamkeitserwartung(t_1), SN = Selbstwirksamkeitserwartung(t_2), 1 = stimme nicht zu, 5 = stimme zu; M = Mittelwert, SD = Standardabweichung.

5.2.3 Flow-Erleben

Bei der Erfassung des Flow-Erlebens (siehe Tabelle 17) mit der Flow-Kurz-Skala (Rheinberg et al., 2019) zeigte sich in der Experimentalgruppe die größte Zustimmung bei Item FL6 (*Ich bin ganz vertieft in das, was ich gerade mache*). Auch Item FL3 (*Ich merke gar nicht, wie die Zeit vergeht*) erhielt eine hohe Zustimmung. Die geringste Zustimmung erhielt in beiden Gruppen Item FL10 (*Ich bin völlig selbstvergessen*). Die Ergebnisse zu den einzelnen Items der Skalen befinden sich in Anhang L. Die Items FL3 und FL6 konnten Rheinberg et al. (2019) bei einer Faktorenanalyse demselben Faktor zuordnen, den sie als *Absorbiertheit* bezeichneten. Zu dieser Komponente zählten auch die Items FL1 und FL10, letzters wies die kleinsten Werte der Zustimmung auf. Hier fielen bereits während der Intervention Verständnisschwierigkeiten seitens der Schüler:innen mit dem Begriff *selbstvergessen* auf, der zu einer geringeren Zustimmung geführt haben könnte. Dies zeigt sich auch in der Kontrollgruppe, bei der Item FL10 ebenfalls die geringste Zustimmung erhielt.

Tabelle 17 Ergebnisse zum Flow-Erleben

Item	Experimentalgruppe		Kontrollgruppe	
	M	SD	M	SD
FL1	3.32	1.20	3.28	1.01
FL3	3.61	1.34	3.52	1.10
FL4	3.30	1.27	3.12	1.18
FL5	3.31	1.29	3.06	1.19
FL6	3.74	1.20	3.29	1.19
FL7	3.37	1.18	3.05	1.22
FL8	3.33	1.11	3.07	1.13
FL9	3.45	1.18	3.27	1.09
FL10	3.10	1.21	2.60	1.09
Gesamt	3.39	0.81	3.14	0.71

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu.

5.2.4 Ergebnisse des computerbezogenen und dem physikbezogenen Selbstkonzept

Die Reliabilitäten McDonalds $\omega = .80$ für das physikbezogene und $\omega = .90$ für das computerbezogene Selbstkonzept waren zufriedenstellend. Bei der Haupterhebung konnten keine Unterschiede zwischen der Experimentalgruppe und der Kontrollgruppe festgestellt werden. Die Mittelwerte und Standardabweichungen der Item-Antworten sind in Tabelle 18 und Tabelle 19 dargestellt. Die Ergebnisse zu den einzelnen Items der Skalen befinden sich in Anhang K.

Bei Betrachtung der Mittelwerte des Antwortverhaltens bei den Items des physikbezogenen Selbstkonzepts fiel auf, dass das Item SKP4 (*Ich glaube, dass mich meine Mitschüler für wirklich gut halten in Physik*) den niedrigsten Mittelwert in beiden Gruppen aufweist. Dieses Item bezieht sich als einziges auf den sozialen Vergleich und stellt damit ein wichtiges Merkmal des Selbstkonzepts dar. Der größte Mittelwert zeigte sich bei Item SKP3 (*In Physik weiß ich schon im Voraus, dass ich nichts verstehe*), was durch die Umkodierung des Items dem geringsten Zuspruch entspricht. Gleichzeitig zeigt dieses Item auch mitunter die größte Standardabweichung. Erklären kann diesen scheinbaren Widerspruch zum Image des Physikunterrichts als unbeliebtes Fach der Zeitpunkt der Erhebung: so wurde direkt nach den Sommerferien erhoben und bislang wurde gemäß Bildungsplan der Physikunterricht größtenteils phänomenbasiert unterrichtet (MKJS, 2015a). Dies sowie die Motivation hinsichtlich der bevorstehenden Intervention kann zu diesem Antwortverhalten geführt haben.

Beim Antwortverhalten des computerbezogenen Selbstkonzepts zeigte sich für Item SKC6 (*Ich habe keine Scheu davor, neue Computerprogramme einfach auszuprobieren*) die höchste Zustimmung. Die Offenheit gegenüber neuen Computerprogrammen könnte möglicherweise durch die Pandemie-

Situation erklärt werden. Während dieser Zeit mussten Schüler:innen mit verschiedenen Videokonferenztools, Lernumgebungen und Plattformen arbeiten. Dies könnte eine positive Wirkung auf ihr computerbezogenes Selbstkonzept gehabt haben. Für die Intervention mit einem Serious Game kann diese Einstellung als positiv bewertet werden, da eine Aufgeschlossenheit angenommen werden konnte. Die geringste Zustimmung fand das Item SKC5 (*Ich habe umfassende Computerkenntnisse*). Dies könnte auf die Formulierung des Items zurückzuführen sein, da es für Schüler:innen schwer einzuschätzen ist, welche Kenntnisse genau gefordert sind. Im Falle des Computers können diese Kenntnisse von der *Hardware*, also dem elektronischen Aufbau des Computers, bis zur *Software*, also sämtlichen Programmen bis zum Betriebssystem, reichen.

Tabelle 18 Ergebnisse des physikbezogenen Selbstkonzepts

Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKP1	3.19	0.98	3.11	1.00
SKP2	3.15	1.21	3.00	1.19
SKP3	3.30	1.36	3.32	1.29
SKP4	2.54	1.16	2.40	1.10
SKP5	3.21	1.26	3.02	1.28
SKP6	3.13	0.99	3.05	1.05
SKP7	2.85	1.32	2.93	1.36
Gesamt	3.05	0.78	2.98	.79

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu. Die Items SKP2, SKP3 und SKP5 wurden umcodiert.

Tabelle 19 Ergebnisse des computerbezogenen Selbstkonzepts

Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKC1	3.44	1.02	3.53	1.03
SKC2	3.55	1.30	3.54	1.20
SKC3	3.15	1.02	3.21	1.07
SKC4	3.00	1.21	3.10	1.13
SKC5	2.91	1.11	2.97	1.04
SKC6	3.74	1.24	3.80	1.19
SKC7	3.22	1.30	3.40	1.23
Gesamt	3.29	0.93	3.36	.85

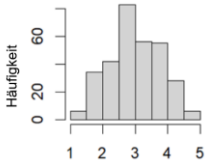
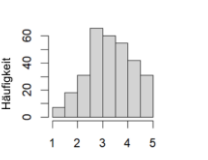
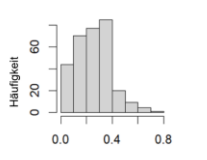
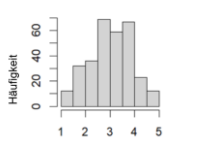
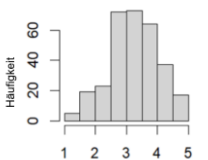
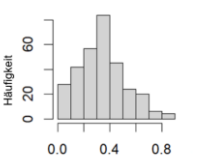
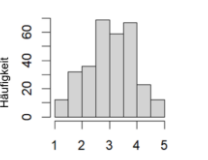
Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu.

5.2.5 Unterschiede nach Gruppen und Geschlechtern

Um die Ähnlichkeit der Voraussetzungen in den Interventionsgruppen zu überprüfen und die Hypothesen der ersten Fragestellung hinsichtlich der Unterschiede zwischen den Gruppen zu validieren, wurden zunächst die Anforderungen für das Verfahren analysiert. Gruppenvergleiche werden typischerweise mit dem t-Test oder Welch's t-Test durchgeführt. Ein signifikanter Nachweis der Alternativhypothese deutet auf einen Unterschied zwischen den Gruppen hin. Es wurden folgende Untersuchungen durchgeführt: Überprüfung auf Ausreißer in Boxplots, Überprüfung auf Normalverteilung mit dem Shapiro-Wilk-Test, Überprüfung auf Kurtosis mit dem Anscombe-Test und auf Schiefe mit dem D'Agostino-Test. Des Weiteren wurde die Homogenität der Varianzen mit dem Levene-Test geprüft. Die Histogramme mit der Verteilung der Antworten sind in Tabelle 20 zu sehen. Der Shapiro-Wilk-Test fiel für fast alle Variablen signifikant aus, was gegen eine Normalverteilung spricht. Jedoch versagt der Shapiro-Wilk-Test bei großen Datenmengen, wie in Kapitel 4.2 beschrieben.

Die optische Prüfung über das Histogramm, der D'Agostino-Test auf Schiefe sowie der Anscombe-Test auf Kurtosis ergaben für die Variablen zufriedenstellende Ergebnisse bezüglich ihrer Normalverteilung. Auch der Levene-Test auf Varianzgleichheit zeigte, mit Ausnahme der Variable *Flow*, gute Ergebnisse. Da die Voraussetzungen für den t-Test nicht erfüllt waren, wurde Welch's t-Test durchgeführt (Kubinger et al., 2009).

Tabelle 20 Prüfung der Ergebnisse auf Normalverteilung

	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
<i>N</i>	310	310	310	310	310	310	310
<i>M</i>	3.02	3.32	.26	3.07	3.28	.34	3.21
<i>SD</i>	.78	.89	.14	.86	.77	.18	.81
Skewness	-.05	-.08	.47	-.15	.07	.46	-.14
Kurtosis	2.39	2.43	3.30	2.49	2.79	2.96	2.74
Shapiro-Wilk	.99***	.99**	.97***	.99**	.99	.98***	.99*
Levene	F(1,308)= .53	F(1,308)= .90	F(1,308)= .40	F(1,308)= .15	F(1,308)= 5.51*	F(1,308)= .81	F(1,308)= .01
Histogramm							

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), CI = Konfidenzintervall, M = Mittelwert, SD = Standardabweichung.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Um zu ermitteln, ob, bedingt durch Schulform oder Erstsprache, zwischen den Schüler:innen unterschiedliche Voraussetzungen existieren, wurden die Ergebnisse dieser Subgruppen im Pretest überprüft. Bei der Gruppierung nach Schulform zeigten sich beim Vergleich der Prädiktorvariablen *computerbezogenes Selbstkonzept* und *Selbstwirksamkeitserwartung* keine signifikanten Unterschiede zwischen den Schüler:innen des Gymnasiums und der Realschule (siehe Tabelle 22). Beim *physikbezogenen Selbstkonzept* sowie dem *Fachwissen(t₁)* und *Fachwissen(t₂)* sind die Differenzen zwischen diesen beiden Gruppen jedoch signifikant: Die Schüler:innen der Realschule verfügten über ein schwach signifikant besseres physikbezogenes Selbstkonzept ($t(236.1) = -2.07, p = .040, d = .24, 95\%-CI [-0.37, -0.01]$). Die Schüler:innen des Gymnasiums erreichten mit schwacher Effektstärke bessere Ergebnisse im Fachwissen Pretest ($t(234) = 2.59, p < .001, d = .30, 95\%-CI [.05, .07]$) sowie mit mittlerer Effektstärke im Posttest ($t(265) = 4.66, p < .001, d = .52, 95\%-CI [.05, .13]$). Wird die Veränderung beim Fachwissen und der Selbstwirksamkeitserwartung von t_1 zu t_2 analysiert, verbesserten sich beide Gruppen signifikant (siehe Tabelle 21 und Tabelle 22). Bemerkenswert ist in diesem Zusammenhang die höhere Effektstärke hinsichtlich der Änderung der Selbstwirksamkeitserwartung bei der Gruppe der Realschüler:innen.

Zwischen Schüler:innen mit Deutsch als Erstsprache und solchen mit einer anderen Erstsprache konnten keine Unterschiede festgestellt werden (siehe Tabelle 22).

Des Weiteren wurde der gesamte Datensatz nach Geschlecht gruppiert und die Ergebnisse verglichen (siehe Tabelle 22). Es konnten, außer für die Variable *Fachwissen(t₁)* und *Fachwissen(t₂)*, signifikante Unterschiede zwischen den Geschlechtern festgestellt werden. So erzielten die Schüler gegenüber den Schülerinnen beim physikbezogenen ($t(306.7) = 5.97, p < .001, d = .68, 95\%-CI [.34, .67]$) und computerbezogenen ($t(306.8) = 8.38, p < .001, d = .95, 95\%-CI [.59, .95]$) Selbstkonzept, beim erlebten Flow während der Intervention ($t(302.6) = 4.59, p < .001, d = .53, 95\%-CI [.23, .59]$), sowie bei der Selbstwirksamkeitserwartung vor der Intervention ($t(293.9) = 5.94, p < .001, d = .67, 95\%-CI [.37, .73]$) und danach ($t(293.3) = 5.85, p < .001, d = .67, 95\%-CI [.34, .69]$) höhere Werte. Zusammengefasst unterscheiden sich die Schüler:innen der beiden Schulformen insbesondere hinsichtlich des bereits vorhandenen Fachwissens in der Elektrizitätslehre und die beiden Geschlechter in Bezug auf die selbstbezogenen Einstellungen.

Tabelle 21 Vergleich der Schulformen zu zwei Zeitpunkten.

	Änderung des Fachwissens			Änderung der Selbstwirksamkeitserwartung		
	Welch's t-Test	95% -CI	Cohens d	Welch-t-Test	95% - CI	Cohens d
RS	$t(108) = -3.06^{**}$	[-.09, -.02]	.29	$t(108) = -3.47^{***}$	[-.37, -.10]	.33
GY	$t(108) = -6.85^{***}$	[-.13, -.07]	.48	$t(200) = -2.17^*$	[-.19, -.01]	.15

Anmerkungen. RS = Realschule, GY = Gymnasium, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Cohens d: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Tabelle 22 Vergleich der Ergebnisse der Subgruppen.

Vergleich nach Schulart							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
Welch's t-Test	$t(236.1) = -2.07^*$	$t(224.2) = -.86$	$t(234) = 2.59^{***}$	$t(228.2) = -.15$	$t(201.2) = 1.05$	$t(265.9) = 4.66^{***}$	$t(190.9) = -1.17$
Cohens d	.24		.30			.52	
95%-CI	[-.37, -.01]		[.01, .07]			[.05, .13]	
RS	3.14	3.29	.23	3.08	3.23	.28	3.29
GY	2.96	3.39	.27	3.06	3.33	.37	3.18
Vergleich nach Erstsprache							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
Welch's t-Test	$t(84) = 1.41$	$t(87.6) = -.37$	$t(82.2) = 1.06$	$t(77.2) = .26$	$t(86.2) = 1.76$	$t(88.1) = 1.85$	$t(73.3) = 1.00$
Deutsch	3.05	3.32	.26	3.08	3.33	.35	3.24
Andere	2.89	3.36	.24	3.04	3.13	.30	3.11
Vergleich nach Geschlecht							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
Welch's t-Test	$t(306.7) = 5.97^{***}$	$t(306.8) = 8.38^{***}$	$t(306.9) = .26$	$t(293.9) = 5.94^{***}$	$t(301.6) = 4.59^{***}$	$t(307.9) = .37$	$t(293.3) = 5.85^{***}$
Cohens d	.68	.95		.67	.53		.67
95%-CI	[.34, .670]	[.59, .95]		[.37, .73]	[.23, .59]		[.34, .69]
m	3.26	3.70	.26	3.34	3.49	.35	3.47
w	2.76	2.93	.25	2.79	3.08	.34	2.95

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), RS = Realschule, GY = Gymnasium, m = männlich, w = weiblich, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Cohens d: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Zur Überprüfung der Hypothesen der ersten Fragestellung wurden die erhobenen Werte der Variablen der Interventionsgruppen miteinander verglichen. Die Ergebnisse sind in Tabelle 24 zu sehen.

Der Vergleich zwischen den beiden Gruppen vor der Intervention im Hinblick auf das computerbezogene und physikbezogene Selbstkonzept sowie die physikbezogene Selbstwirksamkeitserwartung und das Fachwissen ergab keine Unterschiede. Zur Überprüfung der Hypothese 1.1 wurden die Ergebnisse der Interventionsgruppen zum Flow-Erleben verglichen. Dabei konnten signifikant höhere Werte bei der Experimentalgruppe ermittelt werden ($t(301.9) = 3.22$, $p = .01$, $95\%-CI [.11, .47]$), mit einer geringen Effektstärke von $d = .37$. Damit konnte die Hypothese 1.1 verifiziert werden, dass das Serious Game mehr Flow-Erleben ermöglicht als die Materialien der Kontrollgruppe. Dies spricht darüber hinaus für die Vermutung, dass im Rahmen der Pilotierung kein Unterschied zwischen den Interventionsgruppen ermittelt wurde, weil das Spiel zu diesem Zeitpunkt noch fehlerhaft war und Frust hervorrief.

Die Verbesserungen hinsichtlich des Fachwissens und der Selbstwirksamkeitserwartung nach der Intervention wurden mit Welch's t-Test für gepaarte Stichproben überprüft. Die Ergebnisse befinden sich in Tabelle 23. Sowohl die Experimentalgruppe ($t(167) = -4.68$, $p < .001$, $d = .36$, $95\%-CI [-.12, -.05]$) als auch die Kontrollgruppe ($t(141) = -6.12$, $p < .001$, $d = .51$, $95\%-CI [-.12, -.06]$) profitierten signifikant hinsichtlich des Fachwissens. Dies führt zur Ablehnung von Hypothese 1.2, nach der das Spielen des Serious Games einen größeren Zuwachs an Fachwissen zulässt. Welch's t-Test für gepaarte Stichproben ergab in Bezug auf die Verbesserung der Selbstwirksamkeitserwartung für die Experimentalgruppe ein signifikantes Resultat ($t(167) = -3.13$, $p = .002$, $d = .25$, $95\%-CI [-.28, -.06]$), jedoch nicht für die Kontrollgruppe ($t(141) = -2.00$, $p > .05$, $95\%-CI [-.20, .01]$). Somit kann Hypothese 1.3 als bestätigt gelten, dass das Serious Game eine Veränderung der Selbstwirksamkeitserwartung ermöglicht.

Tabelle 23 Vergleich der Interventionsgruppen zu zwei Zeitpunkten.

	Änderung des Fachwissens			Änderung der Selbstwirksamkeitserwartung		
	Welch's t-Test	95% -CI	Cohens d	Welch's t-Test	95% - CI	Cohens d
EX	$t(167) = -4.68^{***}$	$[-.12, -.05]$.36	$t(167) = -3.13^{**}$	$[-.28, -.06]$.25
KG	$t(141) = -6.12^{***}$	$[-.12, -.06]$.51	$t(141) = -2.00$	$[-.20, .01]$	-

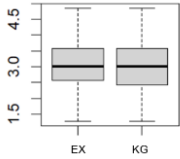
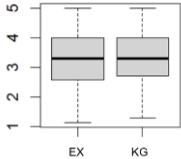
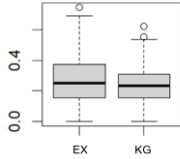
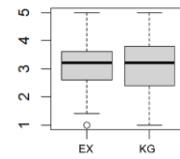
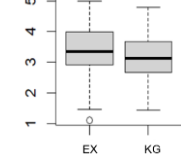
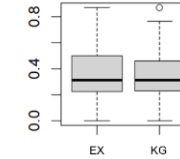
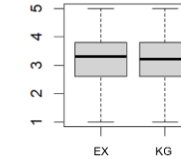
Anmerkungen. EX = Experimentalgruppe, KG = Kontrollgruppe, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$

Cohens d : .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Zur Überprüfung der Hypothesen 1.4 bis 1.9 wurden die beiden Interventionsgruppen jeweils nach Geschlecht gruppiert (Tabelle 25 und Tabelle 26). Die Ergebnisse von Welch's t-Test aus dem Pretest bestätigten sich auch in den einzelnen Interventionsgruppen. So erreichten Schüler bei den Variablen *Selbstkonzepte* und *Selbstwirksamkeitserwartung* höhere Werte als Schülerinnen. Bezüglich des Fachwissens zeigten sich keine Unterschiede. Weiterhin hatten Schüler sowohl in der Experimentalgruppe ($t(165.2) = 2.45$, $p = .02$, $d = .38$, $95\%-CI [.06, .54]$) als auch in der Kontrollgruppe ($t(139.9) = 4.67$, $p < .001$, $d = .78$, $95\%-CI [.30, .74]$) höhere Werte beim erlebten Flow als Schülerinnen.

Tabelle 24 Vergleich der Interventionsgruppen.

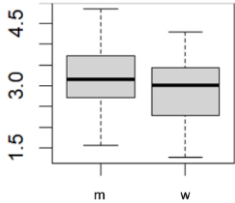
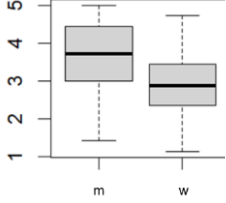
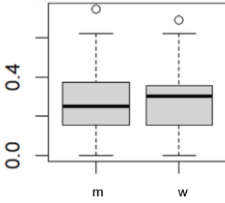
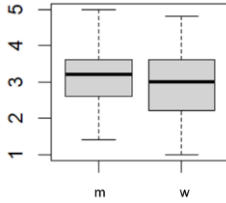
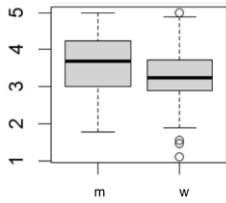
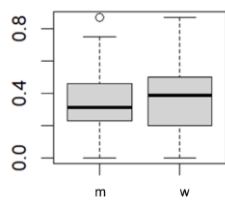
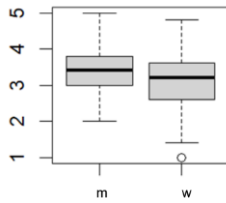
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
Boxplot							
Welch's t-Test	$t(297.1) = .85$	$t(306.2) = -.76$	$t(303.7) = 1.85$	$t(298.9) = .13$	$t(301.9) = 3.22^{**}$	$t(30.6) = 1.02$	$t(298.9) = .13$
Cohens <i>d</i>					.37		
95%-CI					[.11, .47]		
EX	3.06	3.30	.27	3.08	3.43	.35	3.25
KG	2.98	3.36	.24	3.06	3.15	.33	3.16

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), EX = Experimentalgruppe, KG = Kontrollgruppe, CI = Konfidenzintervall

* $p < .05$; ** $p < .01$; *** $p < .001$

Cohens *d*: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Tabelle 25 Geschlechtervergleich in der Experimentalgruppe.

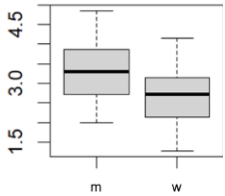
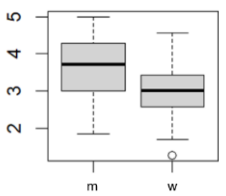
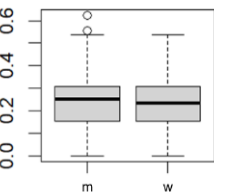
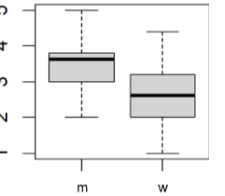
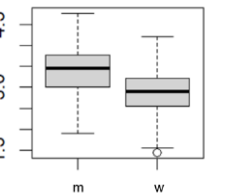
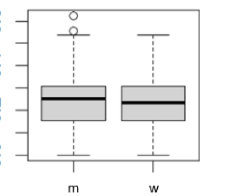
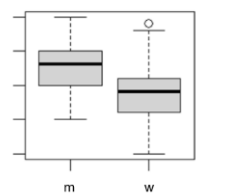
	SKP	SKC	FW _v	SWE _v	Flow	FW _N	SWE _N
Boxplot							
Welch's t-Test	$t(165.2) = 3.22^{***}$	$t(165) = 6.34^{***}$,	$t(165.3) = -.38$	$t(162.8) = 2.37^*$	$t(165.2) = 2.45^*$	$t(165) = -.06$	$t(159.3) = 2.50^{**}$
Cohens <i>d</i>	.50	1.00	-	.36	.38	-	.39
95% - CI	[.15, .61]	[.56, 1.07]		[.05, .57]	[.06, .54]		[.07, .56]
m	3.24	3.69	.27	3.23	3.54	.35	3.40
w	2.86	2.87	.27	2.92	3.24	.35	3.09

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_v = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), m = Schüler, w = Schülerinnen, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$

Cohens *d*: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Tabelle 26 Geschlechtervergleich innerhalb der Kontrollgruppe.

	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
Boxplot							
Welch's t-Test	$t(139.8) = 5.41^{***}$	$t(139.6) = 5.45^{***}$	$t(139.1) = .86$	$t(126.2) = 6.55^{***}$	$t(139.9) = 4.67^{***}$	$t(138.5) = .64$	$t(135.3) = 6.25^{***}$
Cohens <i>d</i>	.91	.91		1.11	.78		1.05
95%-CI	[.42, .90]	[.45, .96]		[.58, 1.09]	[.30, .74]		[.52, 1.00]
m	3.27	3.71	.25	3.46	3.40	.34	3.53
w	2.63	3.03	.23	2.62	2.87	.32	2.77

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂), SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), m = Schüler, w = Schülerinnen, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$

Cohens *d*: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

Bezüglich der Änderung des Fachwissens (Tabelle 27) zeigte sich kein Unterschied zwischen den Geschlechtern: Sowohl die Schüler ($t(84) = -3.58, p < .001, d = .39, 95\%-CI [-.14, -.04]$) als auch die Schülerinnen ($t(82) = -3.04, p = .004, d = .33, 95\%-CI [-.13, -.03]$) der Experimentalgruppe profitierten hinsichtlich der Lernwirksamkeit vom Serious Game.

Auch in der Kontrollgruppe konnten die Schüler ($t(72) = -4.38, p < .001, d = .51, 95\%-CI [-.13, -.05]$) und die Schülerinnen ($t(68) = -4.24, p < .001, d = .51, 95\%-CI [-.13, -.05]$) ihr Fachwissen ausbauen. Bezüglich der Selbstwirksamkeitserwartung war festzustellen, dass in der Experimentalgruppe sowohl die Schüler ($t(85) = -2.28, p = .02, d = .27, 95\%-CI [-.34, -.04]$) als auch die Schülerinnen ($t(82) = -2.19, p = .010, d = .24, 95\%-CI [-.34, -.02]$) profitierten. In der Kontrollgruppe verbesserte sich die Selbstwirksamkeitserwartung weder bei den Schülern noch bei den Schülerinnen.

Tabelle 27 Vergleich der Geschlechter in der Experimental- und Kontrollgruppe zu zwei Zeitpunkten.

	Änderung des Fachwissens			Änderung der Selbstwirksamkeitserwartung		
	Welch's t-Test	95% -CI	Cohens d	Welch's t-Test	95% - CI	Cohens d
EX _M	$t(84) = -3.58^{***}$	[-.14, -.04]	.39	$t(85) = -2.28^{**}$	[-.34, -.04]	.27
EX _W	$t(82) = -3.04^{**}$	[-.13, -.03]	.33	$t(82) = -2.19^*$	[-.34, -.02]	.24
KG _M	$t(72) = -4.38^{***}$	[-.13, -.05]	.51	$t(72) = -.97$	[-.19, .07]	
KG _W	$t(68) = -4.24^{***}$	[-.13, -.05]	.51	$t(67) = -1.87$	[-.29, .01]	

Anmerkungen. EX_M = Schüler der Experimentalgruppe, EX_W = Schülerinnen der Experimentalgruppe, KG_M = Schüler der Kontrollgruppe, KG_W = Schülerinnen der Kontrollgruppe, CI = Konfidenzintervall.

* $p < .05$; ** $p < .01$; *** $p < .001$

Cohens d: .2 bis .4 schwacher Effekt, .5 bis .7 mittlerer Effekt, .8 starker Effekt.

In Tabelle 28 werden alle Hypothesen der ersten Fragestellung aufgelistet und angemerkt, ob sie verifiziert oder verworfen wurden.

Tabelle 28 *Verworfen und verifizierte Hypothesen zur ersten Fragestellung*

Nr.	Alternativhypothese	Verifiziert
H1.1	In der Experimentalgruppe mit einem Serious Game wird mehr Flow erlebt als in der Kontrollgruppe.	✓
H1.2	Die Schüler:innen der Experimentalgruppe profitieren stärker in Bezug auf das Fachwissen als die Schüler:innen der Kontrollgruppe.	✗
H1.3	Die Schüler:innen der Experimentalgruppe profitieren in Bezug auf die Selbstwirksamkeitserwartung stärker als die Schüler:innen der Kontrollgruppe.	✓
H1.4	Innerhalb der Experimentalgruppe gibt es bezüglich des erlebten Flows keine Unterschiede zwischen den Geschlechtern.	✗
H1.5	Innerhalb der Kontrollgruppe erleben die Schüler mehr Flow als die Schülerinnen.	✓
H1.6	Innerhalb der Experimentalgruppe gibt es bezüglich der Verbesserung des Fachwissens keine Unterschiede zwischen den Geschlechtern.	✓
H1.7	Innerhalb der Kontrollgruppe gibt es bezüglich der Verbesserung des Fachwissens keine Unterschiede zwischen den Geschlechtern.	✓
H1.8	Innerhalb der Experimentalgruppe gibt es bezüglich der Selbstwirksamkeitserwartung keine Unterschiede zwischen den Geschlechtern.	✓
H1.9	Innerhalb der Kontrollgruppe profitieren die Schüler bezüglich der Selbstwirksamkeitserwartung stärker als die Schülerinnen.	✗

5.2.6 Ermittlung der Einflussgrößen auf das Erleben von Flow

In der zweiten Fragestellung geht es darum, welche Merkmale (Intervention, Geschlecht, computerbezogenes Selbstkonzept, physikbezogenes Selbstkonzept, Vorwissen und Selbstwirksamkeitserwartung) einen möglichen Einfluss auf das Erleben von Flow haben. Zur Überprüfung der damit zusammenhängenden Hypothesen wurden Regressionsmodelle erstellt.

5.2.6.1 Überprüfung der Voraussetzungen der multiplen Regression

Die Ergebnisse der Voruntersuchung sind in Tabelle 29 dargestellt. Die Daten wurden visuell sowie mithilfe des Breusch-Pagan-Tests auf vorliegende Homoskedastizität geprüft ($p > .05$). Die Normalverteilung der Residuen lag vor, das Histogramm sowie der Quantil-Quantil-Plot zeigten gute optische Ergebnisse und das Resultat des Shapiro-Wilk-Tests war nicht signifikant ($p > .05$). Auch die Prüfung mit dem Durbin-Watson-Test auf Autokorrelation der Residuen war unauffällig. Die Linearität war optisch gegeben, und auch der Rainbow-Test lieferte zufriedenstellende Ergebnisse ($p > .05$). Die Korrelationen der Variablen untereinander wurden nach Pearson (zweiseitig) getestet (Tabelle 30). Die schwache Korrelation des Fachwissens(t_1) mit dem Fachwissen(t_2) ist vermutlich darauf zurückzuführen, dass Schüler:innen mit einem großen Vorwissen von der Lernwirksamkeit nicht mehr profitieren konnten. Andersherum verhält es sich bei Schüler:innen mit wenig Vorwissen. Die Selbstwirksamkeitserwartung(t_1) und die Selbstwirksamkeitserwartung(t_2) sowie das physikbezogene Selbstkonzept und die Selbstwirksamkeitserwartung(t_1) korrelieren stark miteinander. Die Beziehung dieser beiden Prädiktoren bleibt jedoch unterhalb der von Fields empfohlenen Grenze von $r = .80$ (Field et al., 2012). Das Weglassen eines Regressors im Modell soll vermieden werden, da es zu Fehlern führen kann, die als *omitted variable bias* bezeichnet werden und sowohl die unabhängige als auch die abhängige Variable beeinflussen können (Baltès-Götz, 1994).

Tabelle 29 Ergebnisse zur Voruntersuchung der Variablen.

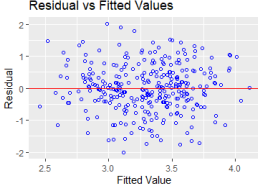
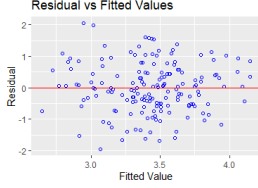
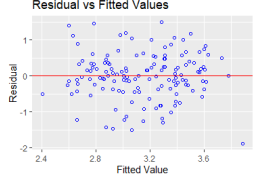
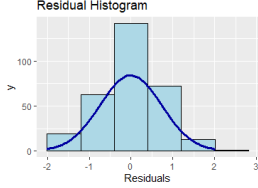
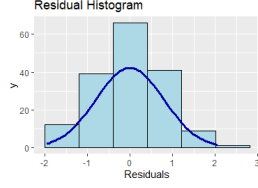
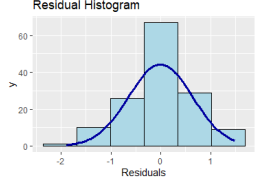
	Alle	Experimentalgruppe	Kontrollgruppe
Homoskedastizität (1) Visuell: Punktwolke (2) Signifikanz mit Breusch-Pagan	 $p = .969$	 $p = .103$	 $p = .847$
Normalverteilung der Residuen: (1) Histogramm, (2) Shapiro-Wilk (3) Schiefe, Signifikanz mit D'Agostino, (4) Kurtosis, Signifikanz mit Anscombe.	 $W = .99, p = .355$ $Schiefe = -.07, p = .566$ $Kurtosis: 2.97, p = .943$	 $W = 1.00, p = .850$ $Schiefe = -.03, p = .835$ $Kurtosis: 2.73, p = .559$	 $W = .97, p = .169$ $Schiefe = -.25, p = .203$ $Kurtosis: 3.27, p = .365$
Linearität: Rainbow Test	$R = 1.29, p = .058$	$R = 1.30, p = .118$	$R = 1.06, p = .399$
Multikollinearität: VIF			
Intervention	1.02	-	-
Geschlecht	1.31	1.26	1.48
Phys. Selbstkonzept	2.13	2.23	2.06
Comp. Selbstkonzept	1.28	1.28	1.28
Phys. SWE(t ₁)	2.08	2.14	2.16
Fachwissen(t ₁)	1.03	1.04	1.01
Autokorrelation der Residuen: Durbin-Watson-Tests	$D-W$ Statistic: 2.05, $p = .720$	$D-W$ Statistic: 1.91, $p = .572$	$D-W$ Statistic: 2.04, $p = .826$

Tabelle 30 Ergebnisse zu den Korrelationen der Variablen.

Korrelationen für die gesamte Erhebung							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
SKP	1						
SKC	.29**	1					
FW _V	.09	-.08	1				
SWE _V	.71**	.24**	.08	1			
Flow	.31**	.25**	-.003	.24**	1		
FW _N	.14*	-.03	.21**	.14*	.17**	1	
SWE _N	.58**	.22**	.06	.68**	.44**	.18**	1
Korrelationen für die Experimentalgruppe							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
SKP	1						
SKC	.23**	1					
FW _V	.13	-.10	1				
SWE _V	.73**	.21**	.07	1			
Flow	.29**	.27**	-.12	.21**	1		
FW _N	.22**	.04	.04	.18*	.14	1	
SWE _N	.50**	.20**	-.01	.62**	.43**	.20*	1
Korrelationen für die Kontrollgruppe							
	SKP	SKC	FW _V	SWE _V	Flow	FW _N	SWE _N
SKP	1						
SKC	.36**	1					
FW _V	.03	-.04	1				
SWE _V	.70**	.28**	.09	1			
Flow	.33**	.27**	.14	.28**	1		
FW _N	.04	-.13	.41**	.10	.20*	1	
SWE _N	.67**	.26**	.13	.76**	.45**	.16	1

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, SWE_V = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), FW_V = Fachwissen(t₁), FW_N = Fachwissen(t₂).

**Die Korrelation ist auf dem Niveau von .01 (zweiseitig) signifikant.

* Die Korrelation ist auf dem Niveau von .05 (zweiseitig) signifikant.

5.2.6.2 Modell 1: Einfluss von Intervention, Geschlecht, Fachwissen, Selbstkonzepten und Selbstwirksamkeitserwartung auf das Flow-Erleben

Mit dem ersten multiplen linearen Regressionsmodell wurde der Einfluss folgender Prädiktoren auf das Flow-Erleben überprüft: Interventionstyp (0 = Experimentalgruppe, 1 = Kontrollgruppe), Geschlecht (0 = Schüler, 1 = Schülerinnen), physikbezogenes sowie computerbezogenes Selbstkonzept, Selbstwirksamkeitserwartung(t_1) und Fachwissen(t_1). Die erklärte Varianz ($R^2 = 16\%$, bzw. korrigiertes $R^2 = 15\%$) durch Zusammenhang der Variablen konnte der F-Test als signifikant ermitteln ($F(6, 303) = 9.95, p < .001$). Die Ergebnisse der Regressionsgleichungen sind in Tabelle 31 zu sehen.

Tabelle 31 Regressionstabelle Flow-Erleben über alle Proban:innen.

	Flow-Erleben				
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Intervention	-.32	-.25	.08	[-.41, -.10]	.003
Geschlecht	-.25	-.19	.09	[-.38, -.01]	.037
SKP	.22	.21	.08	[.06, .36]	.005
SKC	.14	.12	.05	[.05, .47]	.018
SWE _v	.01	.01	.07	[-.27, .28]	.893
FW _v	-.03	-.16	.30	[-1.80, .64]	.583
Konstante	.27		.27	[1.94, 2.99]	< .001

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t_1), SWE_v = Selbstwirksamkeitserwartung(t_1), CI = Konfidenzintervall.
N = 310, $R^2 = 16\%$, korrigiertes $R^2 = 15\%$, Residuen *SE* = .71, $F(6,303) = 9.95, p < .001$.

Signifikante Prädiktoren des Flow-Erlebens sind demnach die Art der Intervention ($B = -.25, p = .003, 95\%-CI [-.41, -.10]$), das Geschlecht ($B = -.19, p = .037, 95\%-CI [-.38, -.01]$), das computerbezogene ($B = .12, p = .018, 95\%-CI [.05, .47]$) und das physikbezogene ($B = .21, p = .005, 95\%-CI [.06, .36]$) Selbstkonzept. Die Ergebnisse dieses Modells zeigen weiterhin, dass die Selbstwirksamkeitserwartung(t_1) und das Fachwissen(t_1) in diesem Zusammenhang keinen Einfluss auf das Flow-Erleben haben ($p > .05$).

5.2.6.3 Modell 2: Einflussfaktoren auf das Flow-Erleben in der Experimentalgruppe

Um zu untersuchen, welche Prädiktoren innerhalb der Experimentalgruppe den größten Einfluss auf das Erleben von Flow haben, wurde ein weiteres Regressionsmodell aufgestellt. Das Ergebnis ist in Tabelle 32 zu sehen.

Tabelle 32 Regressionstabelle Flow-Erleben in der Experimentalgruppe.

	Flow-Erleben				
	β	<i>B</i>	<i>SE</i>	<i>B (95%-CI)</i>	<i>p</i>
Geschlecht	-.10	-.08	.37	[-.34, .18]	.556
SKP	.26	.27	.08	[.05, .50]	.018
SKC	.16	.14	.07	[.01, .29]	.048
SWE _v	-.01	-.01	.10	[-.21, .19]	.916
FW _v	-.14	-.80	.42	[-1.64, .04]	.063
Konstante	.05		.37	[1.66, 3.10]	< .001

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t₁), SWE_v = Selbstwirksamkeitserwartung(t₁), CI = Konfidenzintervall.
N = 168, *R*² = 14%, korrigiertes *R*² = 12%, Residuen *SE* = .76, *F*(5,162) = 5.36, *p* < .001.

Es zeigte sich in Übereinstimmung mit den Hypothesen H2.7 und H2.8, dass das physikbezogene (*B* = .27, *p* = .018, 95%-CI [.05, .50]) und das computerbezogene Selbstkonzept (*B* = .14, *p* = .048, 95%-CI [.01, .29]) einen signifikanten Einfluss haben. Nicht signifikant ist demgegenüber der Einfluss der Selbstwirksamkeitserwartung(t₁), des Fachwissens(t₁) sowie des Geschlechts. Ein unerwartetes Ergebnis ist dabei der negative Koeffizient des Fachwissens(t₁) und der Selbstwirksamkeitserwartung(t₁) mit *p* > .05. Eine mögliche Folgerung könnte sein, dass wenig Vorwissen und eine geringe Selbstwirksamkeitserwartung sich förderlich auf das Flow-Erleben auswirken. Insgesamt kann das Modell einen Anteil von *R*² = 14% (angepasstes *R*² = 12%) der Varianz erklären (*F*(5,162) = 5.36, *p* < .001).

5.2.6.4 Modell 3: Einflussfaktoren auf das Flow-Erleben in der Kontrollgruppe

In der Kontrollgruppe (Tabelle 33) wird ebenfalls das Flow-Erleben durch das physikbezogene Selbstkonzept beeinflusst ($B = .19, p = .049, 95\%-CI [.00, .39]$). Des Weiteren stellt auch das Geschlecht einen signifikanten Prädiktor des Flow-Erlebens dar ($B = -.34, p = .013, 95\%-CI [-.60, -.07]$). Wie erwartet, ist das computerbezogene Selbstkonzept nicht signifikant. Insgesamt erklärt das Regressionsmodell $R^2 = 16\%$ (korrigiertes $R^2 = 14\%$) der Varianz ($F(5,136) = 6.64, p < .001$).

Tabelle 33 Regressionstabelle Flow-Erleben in der Kontrollgruppe.

	Flow				
	β	B	SE	$B (95\%-CI)$	p
Geschlecht	-.47	-.34	.13	[-.60, -.07]	.013
SKP	.21	.19	.10	[.00, .39]	.049
SKC	.12	.10	.07	[-.05, .30]	.187
SWE _v	-.03	-.02	.09	[-.21, .16]	.811
FW _v	.11	.60	.41	[-.22, 1.41]	.149
Konstante	.23		.38	[1.59, 3.09]	< .001

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t_1), SWE_v = Selbstwirksamkeitserwartung(t_1), CI = Konfidenzintervall.

$N = 142, R^2 = 16\%$, korrigiertes $R^2 = 14\%$, Residuen $SE = .66, F(5,136) = 6.64, p < .001$.

Die Gesamtheit der durch Regressionsgleichungen abgelehnten und angenommenen Hypothesen sind in Tabelle 34 aufgeführt.

Tabelle 34 *Verworfenne und verifizierte Hypothesen zur zweiten Fragestellung*

Nr.	Alternativhypothese	Verifiziert
H2.1	Die Interventionsart (Experimentalgruppe oder Kontrollgruppe) hat einen Einfluss auf das Flow-Erleben.	✓
H2.2	Das Vorwissen in der Elektrizitätslehre ist ein Prädiktor für das Flow-Erleben.	✗
H2.3	Das Geschlecht ist in der Experimentalgruppe kein Prädiktor für das Flow-Erleben.	✓
H2.4	Das physikbezogene Selbstkonzept ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.	✓
H2.5	Das computerbezogene Selbstkonzept ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.	✓
H2.6	Die physikbezogene Selbstwirksamkeitserwartung ist in der Experimentalgruppe ein Prädiktor für das Flow-Erleben.	✗
H2.7	Das Vorwissen in der Elektrizitätslehre ist in der Experimentalgruppe kein Prädiktor für das Flow-Erleben.	✗
H2.8	Das Geschlecht ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.	✓
H2.9	Das physikbezogene Selbstkonzept ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.	✓
H2.10	Das computerbezogene Selbstkonzept ist in der Kontrollgruppe kein Prädiktor für das Flow-Erleben.	✓
H2.11	Die physikbezogene Selbstwirksamkeitserwartung ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.	✗
H2.12	Das Vorwissen in der Elektrizitätslehre ist in der Kontrollgruppe ein Prädiktor für das Flow-Erleben.	✗

5.2.7 Wirkungen des Flow-Erlebens

Um zu untersuchen, inwiefern das computerbezogene und das physikbezogene Selbstkonzept, das Fachwissen(t_1) und die Selbstwirksamkeitserwartung(t_1) in Zusammenhang mit dem Flow-Erleben das Fachwissen(t_2) und die Selbstwirksamkeitserwartung(t_2) vorhersagen, wurde ein Strukturgleichungsmodell für die beiden Interventionsgruppen erstellt. Die Selbstkonzepte, sowie die Selbstwirksamkeitserwartung vor und nach der Intervention wurden als manifeste Variablen in das Modell eingefügt. Aufgrund der geringen Stichprobe wurde das Item-Parceling auf die Variable Flow-Erleben angewendet, einerseits um die Einflüsse der Fehlervarianz durch die Skalen zu vermeiden (siehe Abschnitt 5.2.1), des Weiteren um die Anzahl der zu schätzenden Parameter zu verringern. Dabei wurden die Items so auf die Parcels verteilt, dass jedes Parcel eine ähnliche Gesamtvarianz hat, Parcel 1 enthielt die Items FL6, FL4, FL8, Parcel 2 die Items FL5, FL7, FL3 und Parcel 3 = FL9, FL1, FL2. Das konfirmatorische Faktorenmodell wies sehr gute Werte auf (CFI & TFI = 1.0, RMSEA = .00, SRMR = .00), die Ladungen auf den Faktor Flow sind signifikant und moderat bis hoch (.61 bis .78 in standardisierter Form). Die Varianzen sind alle signifikant und positiv, was darauf hinweist, dass die Parcels nicht alle Varianz in den Items erklären. Die Varianz von Flow ist signifikant und gleich 1 in standardisierter Form, was darauf hindeutet, dass Flow eine signifikante und angemessene Menge an Varianz erklärt. Die R^2 -Weerte sind alle moderat bis hoch (.54 bis .69), was darauf hindeutet, dass eine beträchtliche Menge der Varianz in den Parcels durch Flow erklärt wird.

Durch Bildung des Mittelwertes wurden die Skalen der Selbstkonzepte und der physikbezogenen Selbstwirksamkeitserwartung als manifeste Variablen im Strukturgleichungsmodell berücksichtigt. Auch die Ergebnisse des Fachwissenstest(t_1) und (t_2) wurden als manifeste Variable eingeschlossen, da aufgrund der Einteilung in A- und B-Tests die jeweiligen Subgruppen zu klein gewesen wären. Da es sich bei dem Ergebnis des Fachwissenstests zudem um eine direkt beobachtbare Variable handelt, geht hierdurch auch keine Varianz verloren. Ein weiterer Vorteil dieses Verfahrens stellt die Anschlussfähigkeit an die Ergebnisse der multiplen linearen Regressionen im vorigen Abschnitt, sowie an die Verwendung der PROCESS Methoden dar, für die ebenfalls die Mittelwerte gebildet werden. Einen Vorteil stellt in lavaan die Berücksichtigung von Kovarianzen dar, so wurde für das physikbezogene Selbstkonzept und die physikbezogene Selbstwirksamkeitserwartung(t_1) die gemeinsame Kovarianz angefordert.

5.2.7.1 Strukturgleichungsmodell der Experimentalgruppe

Der Fit des Modells zu den Daten der Experimentalgruppe ist insgesamt akzeptabel ($\chi^2 = 10.38$, $p > .05$, $df = 5$, $\chi^2/df = 2.07$ (akzeptabel), $SRMR = .05$ (gut), $RMSEA = .08$ (akzeptabel), $CFI = .98$ (gut), $TLI = .95$ (akzeptabel)).

Das Latente Messmodell für Flow sind in *Tabelle 35* und die Ergebnisse der Regressionen sind in *Tabelle 36* zu sehen, das Pfaddiagramm in *Abbildung 22*. Die berücksichtigte Kovarianz zwischen physikbezogenem Selbstkonzept und physikbezogener Selbstwirksamkeitserwartung war signifikant ($B = .45$, $p < .001$, $95\%-CI [.33, .57]$). Signifikante Zusammenhänge bestehen zwischen dem Geschlecht und den beiden Selbstkonzepten sowie der physikbezogenen Selbstwirksamkeitserwartung. Kein Zusammenhang konnte zwischen Geschlecht und Fachwissen(t_1) ermittelt werden, sodass die Ergebnisse diesbezüglich der ersten und zweiten Fragestellung auch in diesem Modell bestätigt werden.

Ebenfalls reproduzieren lassen sich im Strukturgleichungsmodell die Einflüsse der Selbstkonzepte auf das Flow-Erleben. Diese waren im multiplen Regressionsmodell im vorigen Abschnitt signifikante

Prädiktoren des Flow-Erlebens, während das Fachwissen(t_1) keinen Einfluss zeigte. Im gemeinsamen Strukturgleichungsmodell aller Variablen lassen sich auch diese Ergebnisse bestätigen.

Keine der Variablen war jedoch ein Prädiktor für das Fachwissen(t_2), während sich auf die Selbstwirksamkeitserwartung(t_2) zwei hoch signifikante Pfade zeigen: des Flow-Erlebens ($B = .35$, $p < .001$, 95%-CI [.19, .42]) sowie der Selbstwirksamkeitserwartung(t_1) ($B = .52$, $p < .001$, 95%-CI [.38, .67]). Bemerkenswert ist, dass das Geschlecht keinen Effekt auf die Selbstwirksamkeitserwartung(t_2) zeigt, die es jedoch noch auf die Selbstwirksamkeitserwartung(t_1) vor der Intervention hatte. Ein Zusammenhang der Variablen des Pretests mit jenen des Posttests kann möglicherweise durch eine Moderator- und Mediatoranalyse des Flow-Erlebens erklärt werden. Dies wird in Abschnitt 5.2.7.3 untersucht.

Tabelle 35. Messmodell des Flow-Erlebens in der Experimentalgruppe

	Flow-Erleben				
	β	B	SE	$B(95\%-CI)$	p
Parcel 1	.84	1	-	-	-
Parcel 2	.77	.81	.08	[.66, .98]	<.001
Parcel 3	.80	.88	.08	[.07, 1.10]	<.001

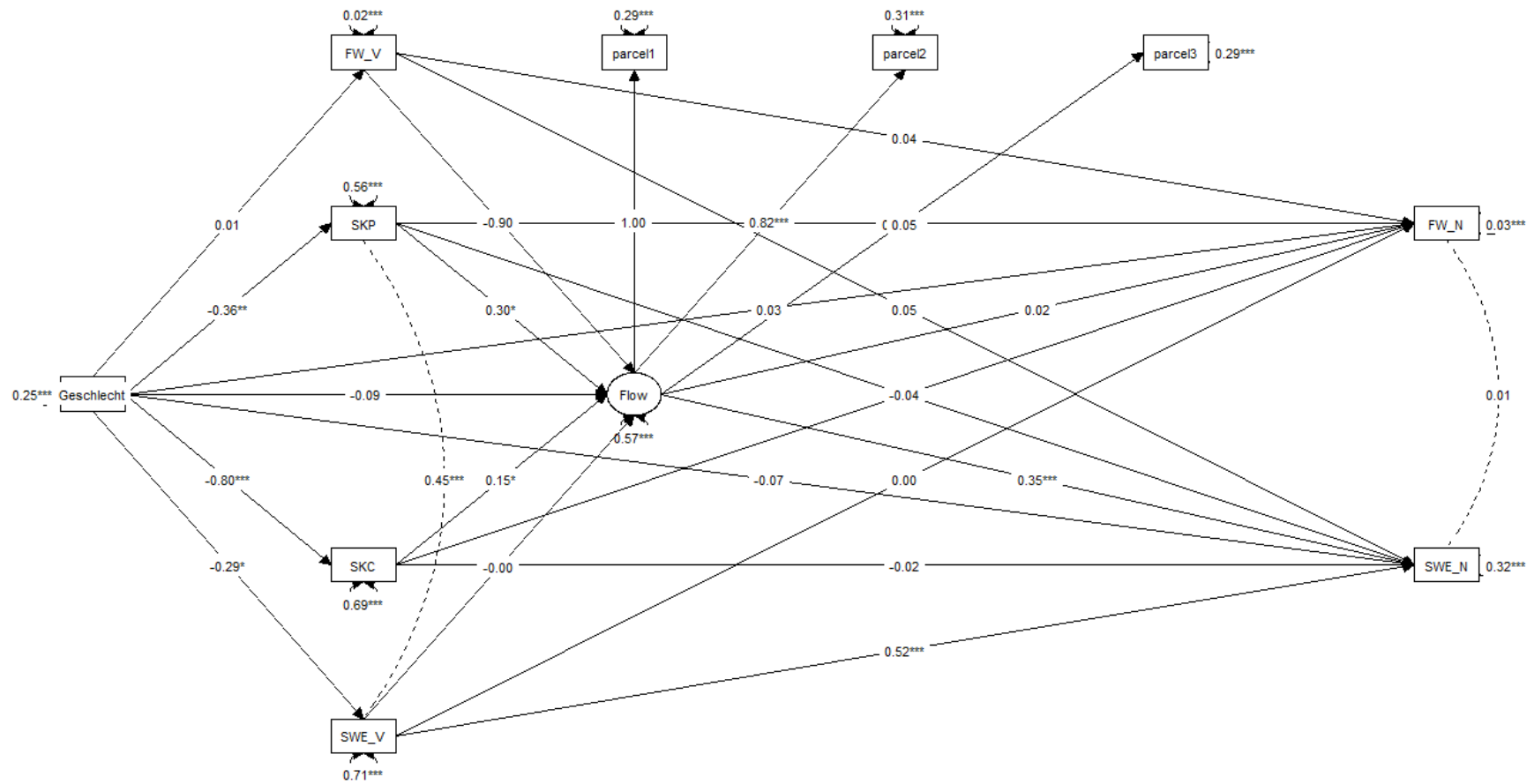
Tabelle 36 Regressionstabelle des Strukturgleichungsmodells für die Experimentalgruppe.

Physikbezogenes Selbstkonzept					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.24	-.36	.11	[-.58, -.14]	.002
Computerbezogenes Selbstkonzept					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
G	-.43	-.80	.13	[-1.00, -.50]	< .001
Fachwissen(t ₁)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	.03	.01	.02	[-.03, .05]	.661
Selbstwirksamkeitserwartung(t ₁)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.17	-.29	.13	[-.55, -.03]	.025
Flow					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.06	-.09	.15	[-.35, .17]	.495
SKP	.28	.29	.12	[.06, .49]	.014
SKC	.17	.15	.08	[.00, .27]	.045
FW _v	-.15	-.90	.45	[-1.61, .00]	.051
SWE _v	<-.01	<-.01	.11	[-.20, .18]	.931
Fachwissen(t ₂)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	.08	.03	.03	[-.03, .09]	.384
Flow	.11	.02	.02	[-.01, .06]	.235
SKP	.20	.05	.03	[.00, .10]	.051
SKC	.01	.02	.02	[-.03, .03]	.954
FW _v	.03	.10	.09	[-.16, .22]	.730
SWE _v	< .01	.02	.02	[-.04, .05]	.985
Selbstwirksamkeitserwartung(t ₂)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.04	-.07	.10	[-.27, .13]	.483
Flow	.36	.35	.06	[.19, .42]	< .001
SKP	-.04	-.04	.09	[-.19, .15]	.840
SKC	-.03	-.02	.05	[-.12, .09]	.836
FW _v	.01	-.06	.32	[-.64, .62]	.970
SWE _v	.56	.52	.08	[.38, .67]	< .001
Kovarianzen					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
SKP und SWE _v	.72	.45	.06	[.33, .57]	<.001
FW _N und SWE _N	.09	.01	.01	[-.01, .03]	.195

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t₁), SWE_v = Selbstwirksamkeitserwartung(t₁), FW_N = Fachwissen(t₂), SWE_N = Selbstwirksamkeitserwartung(t₂), CI = Konfidenzintervall.

N = 168, $\chi^2 = 10.38$, *p* > .05, *df* = 5, $\chi^2/df = 2.07$, SRMR = .05, RMSEA = .08, CFI = .98, TLI = .95.

Abbildung 22 Pfaddiagramm der Experimentalgruppe.



Anmerkungen. Die Pfade sind unstandardisiert., SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), SWE_V = Selbstwirksamkeitserwartung(t₁), FW_N = Fachwissen(t₂), SWE_N = Selbstwirksamkeitserwartung(t₂). *p < .05; **p < .01; ***p < .001. Durchgezogene Linien beschreiben Pfade, gestrichelte Linien Kovarianzen

5.2.7.2 Strukturgleichungsmodell der Kontrollgruppe

Das Modell der Kontrollgruppe passte insgesamt, gut zu den erhobenen Daten ($\chi^2 = 8.91$, $p > .05$, $df = 5$, $\chi^2/df = 1.78$ (gut), $SRMR = .04$ (gut), $RMSEA = .07$ (akzeptabel), $CFI = .99$ (gut), $TLI = .94$ (akzeptabel)), die Ergebnisse des Messmodells für Flow-Erleben sind in *Tabelle 37* und die Ergebnisse der Regressionen sind in *Tabelle 38* abgebildet, das Pfaddiagramm in *Abbildung 23*. Auch hier ließen sich die signifikanten Unterschiede zwischen den Geschlechtern hinsichtlich der Selbstkonzepte und der physikbezogenen Selbstwirksamkeitserwartung erkennen. Ebenfalls konnte hier kein Einfluss des Geschlechts auf das Fachwissen(t_1) ermittelt werden.

Auch der signifikante Einfluss des Geschlechts auf das Flow-Erleben ließ sich hier erkennen und wie in der zweiten Fragestellung ist das physikbezogene Selbstkonzept ein signifikanter Prädiktor des Flow-Erlebens.

Bemerkenswert ist, dass das Flow-Erleben ($B = .05$, $p = .034$, $95\%-CI [.00, .09]$) sowie das Fachwissen(t_1) ($B = .48$, $p < .001$, $95\%-CI [.30, .69]$) in der Kontrollgruppe einen signifikanten Einfluss auf das Fachwissen(t_2) aufweist. Dieser Zusammenhang soll in einer Moderationsanalyse im nächsten Abschnitt weiter untersucht werden. Überraschend ist hier darüber hinaus die signifikante Wirkung des computerbezogenen Selbstkonzepts ($B = -.04$, $p = .032$, $95\%-CI [-.07, .00]$) auf das Fachwissen(t_2). Die Selbstwirksamkeitserwartung(t_2) wird, ähnlich dem Modell der Experimentalgruppe, auch hier durch das Flow-Erleben ($B = .26$, $p < .001$, $95\%-CI [.13, .38]$) und die Selbstwirksamkeitserwartung(t_1) ($B = .49$, $p < .001$, $95\%-CI [.13, .38]$) vorhergesagt. Zusätzlich zeigt auch das physikbezogene Selbstkonzept hier einen signifikanten Einfluss. Das Geschlecht zeigt hier keinen signifikanten Einfluss. Weiteren Aufschluss über die Wirkungen der Prädiktoren und des Flow-Erlebens auf das Fachwissen(t_2) und die Selbstwirksamkeitserwartung(t_2) sollen im nächsten Abschnitt ermittelt werden.

Tabelle 37: Messmodell des Flow-Erlebens in der Kontrollgruppe

	Flow-Erleben				
	β	B	SE	$B(95\%-CI)$	p
Parcel 1	.86	1	-	-	-
Parcel 2	.83	.91	.10	[.07, 1.10]	<.001
Parcel 3	.61	.55	.08	[.40, .70]	<.001

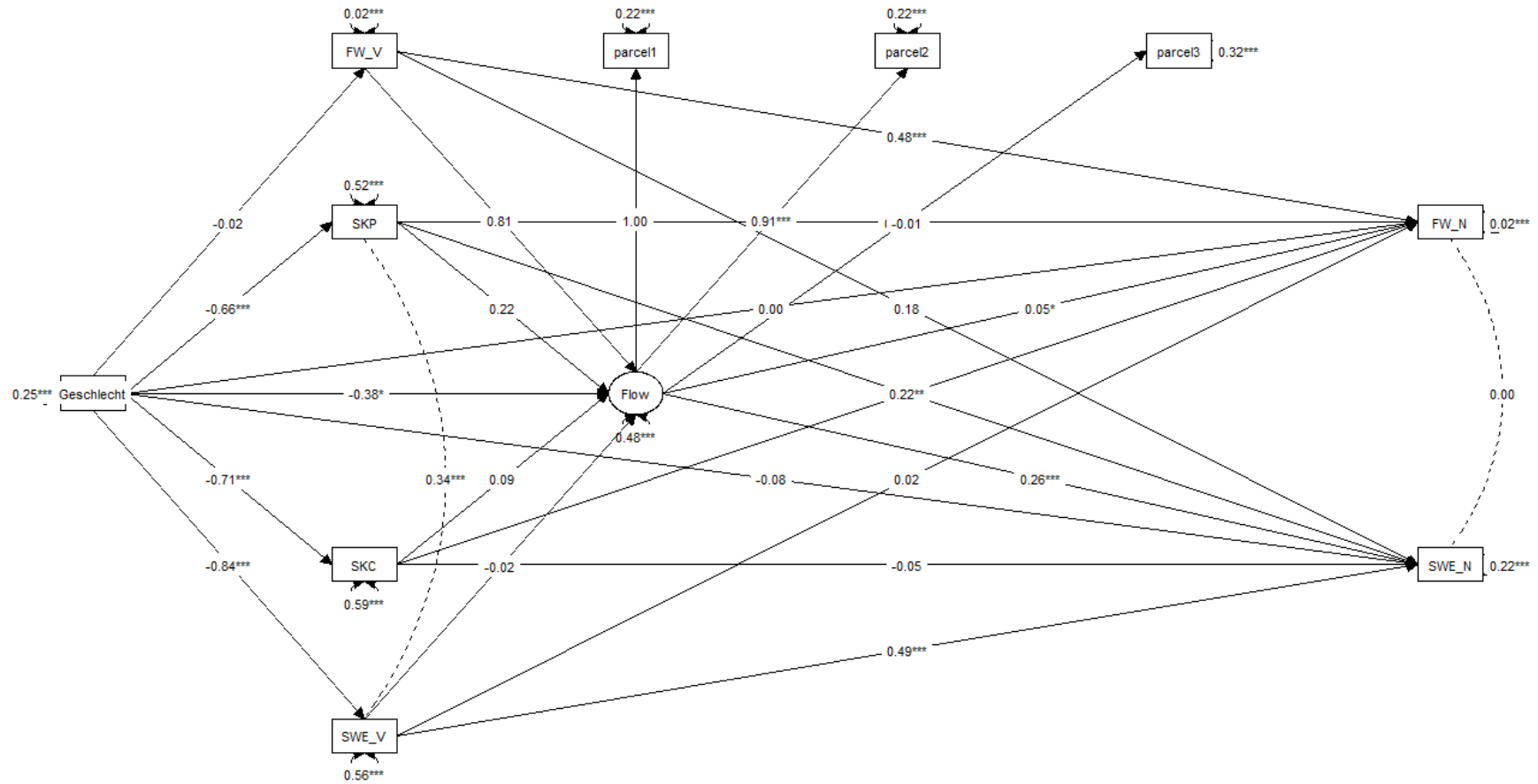
Tabelle 38 Regressionstabelle des Strukturgleichungsmodells für die Kontrollgruppe.

Physikbezogenes Selbstkonzept					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.42	-.66	.12	[-.89, -.42.]	<.001
Computerbezogenes Selbstkonzept					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.42	-.71	.13	[-.96, -.45]	<.001
Fachwissen(t ₁)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.07	.02	.02	[-.06, .01]	.386
Selbstwirksamkeitserwartung(t ₁)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.49	-.84	.14	[-1.14, -.58]	<.001
Flow					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.24	-.38	.16	[-.59, -.07]	.012
SKP	.23	.22	.09	[.00, .38]	.044
SKC	.10	.09	.07	[-.04, .23]	.184
FW _v	.14	-.81	.40	[-.16, 1.41]	.120
SWE _v	-.02	-.02	.09	[-.20, .15]	.785
Fachwissen(t ₂)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	<.01	<.01	.04	[-.07, .06]	.956
Flow	.22	.05	.06	[.00, .09]	.034
SKP	-.05	-.01	.03	[-.05, .04]	.726
SKC	-.18	-.04	.02	[-.07, .00]	.032
FW _v	.37	.48	.10	[.30, .69]	<.001
SWE _v	.09	.02	.03	[-.02, .06]	.384
Selbstwirksamkeitserwartung(t ₂)					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
Geschlecht	-.05	-.08	.10	[-.29, .10]	.349
Flow	.25	.26	.06	[.13, .38]	<.001
SKP	.22	.22	.07	[.09, .37]	.001
SKC	-.05	-.05	.05	[-.15, .05]	.360
FW _v	.03	.18	.30	[-.36, .82]	.446
SWE _v	.52	.49	.07	[.36, .63]	<.001
Kovarianzen					
	β	<i>B</i>	<i>SE</i>	<i>B(95%-CI)</i>	<i>p</i>
SKP und SWE _v	.62	.34	.05	[.23, .44]	<.001
FW _N und SWE _N	.03	<.01	.01	[-.01, .02]	.527

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t₁), SWE_v = Selbstwirksamkeitserwartung(t₁), FW_N = Fachwissen(t₂), SWE_N = Selbstwirksamkeitserwartung(t₂), CI = Konfidenzintervall.

$N = 142$, $\chi^2 = 8.91$, $p > .05$, $df = 5$, $\chi^2/df = 1.78$, $SRMR = .04$, $RMSEA = .07$, $CFI = .99$, $TLI = .94$;

Abbildung 23 Pfaddiagramm der Kontrollgruppe.



Anmerkungen. Die Pfade sind unstandardisiert. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_V = Fachwissen(t₁), SWE_V = Selbstwirksamkeitserwartung(t₁), FW_N = Fachwissen(t₂), SN = Selbstwirksamkeitserwartung(t₂). *p < .05; **p < .01; ***p < .001. Durchgezogene Linien beschreiben Pfade, gestrichelte Linien Kovarianzen

5.2.7.3 Ermittlung der Rolle des Flow-Erlebens bei der Veränderung von Fachwissen und Selbstwirksamkeitserwartung zwischen zwei Zeitpunkten

Zur Überprüfung der Hypothese, wonach das Flow-Erleben als Moderator für die Änderung des Fachwissens und der Selbstwirksamkeitserwartung fungiert, wurde eine Moderationsanalyse mit PROCESS (Hayes, 2017) durchgeführt. Untersucht wurde die Moderatorrolle des Flow-Erlebens bei der Wirkung des Fachwissens(t_1) auf das Fachwissen(t_2) respektive der Selbstwirksamkeitserwartung(t_1) auf die Selbstwirksamkeitserwartung(t_2). Die Variablen Geschlecht, physikbezogenes und computerbezogenes Selbstkonzept sowie ergänzend Fachwissen(t_1) bzw. Selbstwirksamkeitserwartung(t_1) wurden als Kovariaten in diesem Modell berücksichtigt.

Experimentalgruppe

Für die Experimentalgruppe zeigte sich bei der Moderation der Wirkung des Fachwissens(t_1) auf das Fachwissen(t_2) weder der Interaktionsterm noch das Gesamtmodell als signifikant.

Die Mediationsanalyse ergab, dass der Gesamteffekt ($B = .52$, $p < .001$) von Selbstwirksamkeitserwartung(t_1) auf Selbstwirksamkeitserwartung(t_2) sowie der direkte Effekt ($B = .52$, $p < .001$) signifikant waren. Dies deutet darauf hin, dass es einen signifikanten direkten Effekt der Selbstwirksamkeitserwartung(t_1) auf die Selbstwirksamkeitserwartung(t_2) gibt, wenn man den Mediator Flow-Erleben kontrolliert. Der indirekte Effekt durch den Mediator Flow-Erleben war nicht signifikant ($B = -0.01$, $p > 0.05$). Dies deutet darauf hin, dass Flow-Erleben nicht als Mediator in der Beziehung zwischen der Selbstwirksamkeitserwartung(t_1) und der Selbstwirksamkeitserwartung(t_2) wirkt.

Kontrollgruppe

In der Kontrollgruppe konnte ein signifikanter Interaktionseffekt ($B = .29$, $p = .043$, 95%-CI [.01, .57]) der ebenfalls signifikanten Variablen Fachwissen(t_1) ($B = .49$, $p < .001$, 95%-CI [.30, .70]) und Flow-Erleben ($B = .04$, $p = .036$, 95%-CI [.01, .09]) ermittelt werden. Das Gesamtmodell ($F(7, 134) = 6.10$, $p < .001$) konnte eine große Varianz von $R^2 = 24\%$ bzw. korrigiertes $R^2 = 20\%$ erklären. Alle Ergebnisse sind in Tabelle 39 zu sehen.

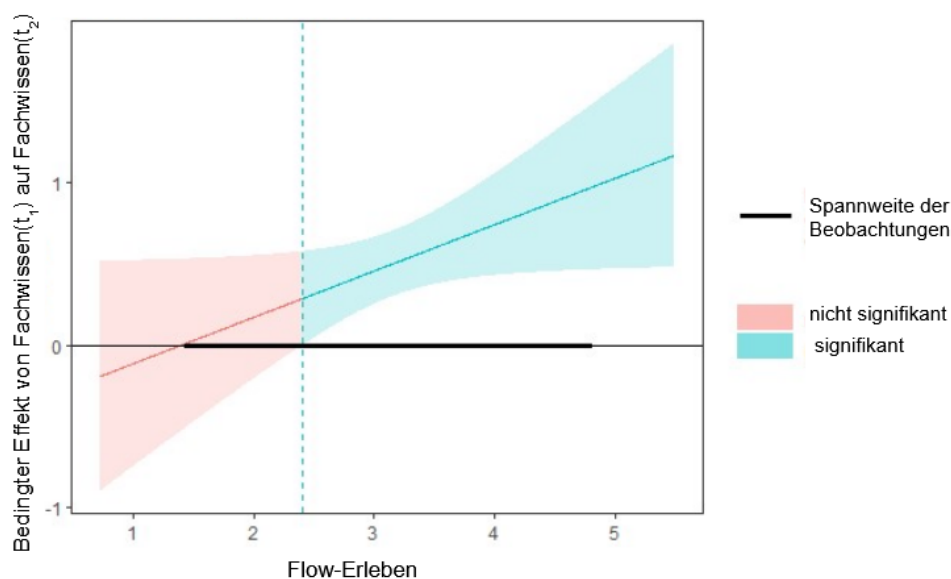
Tabelle 39 Moderatoranalyse von Flow und Fachwissen.

	Fachwissen(t ₂)				
	β	<i>B</i>	<i>SE</i>	<i>B</i> (95%-CI)	<i>p</i>
Flow	2.12	.04	.02	[.01, .09]	.036
FW _v	4.91	.49	.10	[.30, .70]	< .001
Interaktion	2.01	.29	.14	[.01, .57]	.043
Kovariaten					
Geschlecht	-.03	-.01	.03	[.07, .06]	.975
SKP	-.03	-.01	.02	[-.05, .05]	.976
SKC	-1.94	.01	.01	[-.07, .01]	.055
SWE _v	.79	.02	.02	[-.03, .06]	.446
Konstante	4.31	.39	.09	[.21, .57]	< .001

Anmerkungen. SKP = physikbezogenes Selbstkonzept, SKC = computerbezogenes Selbstkonzept, FW_v = Fachwissen(t₁), SWE_v = Selbstwirksamkeitserwartung(t₁), SWE_N = Selbstwirksamkeitserwartung(t₂), CI = Konfidenzintervall. *N* = 142, *F*(7,134) = 6.10, *p* < .001, *R*² = 24%, korrigiertes *R*² = 20%, Interaktionsterm *R*² = 3%.

Im *Jonson-Neyman-Plot* (Abbildung 24) kann der Bereich der Interaktion abgelesen werden. Die X-Achse repräsentiert die Zustimmung zu den Items der Flow-Kurz-Skala. Zu erkennen ist hier an der schwarzen Linie, dass die erfassten Mittelwerte der Schüler:innen im Intervall zwischen 1.44 (geringer erlebter Flow) bis 4.78 (höchste Werte beim erlebten Flow). Die obere und untere Begrenzung des Bereichs stellen die obere und untere Grenze des 95% Konfidenzintervalls dar. Auf der Y-Achse ist der Effekt der Prädiktorvariable Fachwissens(t₁) auf das Fachwissen(t₂) aufgetragen. Dem Jonson-Neyman-Intervall ist zu entnehmen, dass ab einem Wert von 2.41 beim Flow-Erleben der bedingte Effekt von Fachwissen(t₁) auf Fachwissen(t₂) signifikant ist (*p* < .05). Dies lässt die Vermutung zu, dass Flow-Erleben in der Kontrollgruppe die Veränderung des Fachwissens moderiert.

Abbildung 24 Johnson-Neyman-Plot des Moderatoreffekts von Flow auf das Fachwissen.



Hinsichtlich der Untersuchung der Mediatorrolle des Flow-Erlebens wurde, wie in der Experimentalgruppe, der Gesamteffekt ($B = .50, p = <.001$) und der indirekte Effekt ($B = .50, p < .001$) waren Signifikant, ohne dass der Mediationseffekt des Flow-Erlebens ($B = -.01, p > .05$) signifikant wurde.

Welche Hypothesen der dritten Fragestellung angenommen und welche verworfen wurden, ist Tabelle 40 zu entnehmen. Die Implikationen dieser Ergebnisse sollen im nächsten Kapitel besprochen werden.

Tabelle 40 *Verworfenne und verifizierte Hypothesen zur dritten Fragestellung*

Nr.	Alternativhypothese	verifiziert
H3.1	Das Flow-Erleben ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.	✗
H3.2	Das Flow-Erleben ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.	✓
H3.3	Das Flow-Erleben ist ein Moderator für die Veränderung des Fachwissens in der Experimentalgruppe.	✗
H3.4	Das Flow-Erleben ist ein Mediator für die Veränderung der Selbstwirksamkeitserwartung in der Experimentalgruppe.	✗
H3.5	Das physikbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.	✗
H3.6	Das computerbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Experimentalgruppe.	✗
H3.7	Das physikbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.	✗
H3.8	Das computerbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Experimentalgruppe.	✗
H3.9	Das Flow-Erleben ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.	✓
H3.10	Das Flow-Erleben ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.	✓
H3.11	Das Flow-Erleben ist ein Moderator für die Veränderung des Fachwissens in der Kontrollgruppe.	✓
H3.12	Das Flow-Erleben ist ein Mediator für die Veränderung der Selbstwirksamkeitserwartung in der Kontrollgruppe.	✗
H3.13	Das physikbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.	✗
H3.14	Das computerbezogene Selbstkonzept ist ein Prädiktor für das Fachwissen in der Kontrollgruppe.	✓
H3.15	Das physikbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.	✓
H3.16	Das computerbezogene Selbstkonzept ist ein Prädiktor für die Selbstwirksamkeitserwartung in der Kontrollgruppe.	✗

6 Diskussion der Ergebnisse

Um zu untersuchen, welche Wirkungen ein Serious Game bezüglich Lernens und Selbstwirksamkeitserwartung hat, wurde eine Studie im Pre-Post-Design mit Kontrollgruppe durchgeführt. Dabei wurden im Pretest die Variablen Geschlecht, computerbezogenes und physikbezogenes Selbstkonzept, Selbstwirksamkeitserwartung und Fachwissen erhoben. Anschließend wurden die Schüler:innen zufällig der Experimental- oder Kontrollgruppe zugeteilt und während der Intervention das Flow-Erleben gemessen. In der Experimentalgruppe fand dies im Serious Game durch Pop-Ups statt, in denen Fragen beantwortet werden mussten, in der Kontrollgruppe in Form von Papierbögen. Nach der Intervention wurde erneut das Fachwissen und die Selbstwirksamkeitserwartung erhoben.

Im Folgenden werden die Ergebnisse der Erhebung diskutiert. Zunächst werden die Ergebnisse in der zeitlichen Reihenfolge ihrer Erhebung zusammengefasst und im Zusammenhang mit der theoretischen Herleitung und des nomologischen Netzwerks interpretiert. Die Buchstaben in Klammern beziehen sich dabei auf die Pfade der nomologischen Netzwerke in Abbildung 25 und Abbildung 26. Anschließend liegt der Fokus auf den Limitierungen und ein Ausblick wird formuliert.

6.1 Pretest

Zunächst konnten Unterschiede zwischen den Schulen bezüglich des vorhandenen Fachwissens(t_1) festgestellt werden. Gründe hierfür können bspw. in Unterschieden bezüglich der kognitiven Aktivierung der Schüler:innen (Szogs et al., 2016), der Förderung des Interesses und der Unterrichtsqualität liegen (Welberg et al., 2021). Auch die höhere Effektstärke der Änderung der Selbstwirksamkeitserwartung bei der Gruppe der Realschüler:innen könnte möglicherweise so begründet werden: So haben sich die Schüler:innen durch die Intervention als wirksam erlebt, wie es sonst der Physikunterricht an der Realschule nicht ermöglicht.

Überraschenderweise zeigten die Ergebnisse, dass es keinen Unterschied im Fachwissen zwischen den Schüler:innen, die angaben, dass Deutsch ihre Erstsprache ist, und den Schüler:innen mit einer anderen Erstsprache gab. Insbesondere im Physikunterricht ist das Erlernen der Fachsprache ein grundlegender Aspekt (Rincke & Leisen, 2015), von dem vermutet wurde, dass er sich auf die Punktzahl im Fachwissenstest auswirkt. Darüber hinaus können Kinder mit sprachlichen Unsicherheiten negative Erfahrungen im Unterricht gemacht haben, die ihr Selbstkonzept und ihre Selbstwirksamkeitserwartung beeinflussen. Es ist jedoch zu beachten, dass nur ein kleiner Anteil der Kinder angab, dass Deutsch nicht ihre Erstsprache ist. Dies könnte zu statistischen Unsicherheiten führen. Darüber hinaus besteht die Möglichkeit, dass die betreffenden Kinder so fließend Deutsch sprechen, dass die Sprache für sie kein Hindernis darstellt und daher keine Defizite erkennbar sind. Auch konnten keine Unterschiede zwischen den Geschlechtern (a) hinsichtlich des vorhandenen Fachwissens festgestellt werden, dies steht in Einklang mit den Ergebnissen großer Leistungsstudien wie PISA (OECD, 2016). Jedoch ist zu erkennen, dass insbesondere die Fragen zu allgemeinen Verständnis aus dem Kompetenzbereich 1 richtig beantwortet wurden (Itemformulierungen siehe Anhang D). In beiden Gruppen bereiteten die Fragen zu Parallelschaltungen (B9 und B10) Schwierigkeiten, was auch im Zusammenhang mit der Fehlvorstellung *Mangelnde Unterscheidung zwischen Reihen- und Parallelschaltung* zu sehen ist (siehe Kapitel 4.1.1). Mögliche Gründe hierfür liegen einerseits in der größeren Schwierigkeit des Themas sowie im geforderten Kompetenzbereich II, aber auch die Müdigkeit durch die Dauer der Intervention kann eine Rolle gespielt haben.

Bezüglich der selbstbezogenen Einstellungen konnten Unterschiede zwischen den Geschlechtern (a) identifiziert werden. So erzielten Schüler in der Befragung signifikant höhere Werte für das computerbezogene sowie das physikbezogene Selbstkonzept und für die physikbezogene Selbstwirksamkeitserwartung als Schülerinnen. Dies deckt sich mit den Ergebnissen von Studien wie PISA (OECD, 2009, 2016). Demnach haben Schülerinnen trotz gleicher Physikleistungen wie Schüler geringe physikbezogene Selbstbilder (Solga & Pfahl, 2009).

Die signifikanten Unterschiede zwischen den Geschlechtern beim computerbezogenen Selbstkonzept stimmen ebenfalls mit den Befunden vorangegangener Studien überein (Eickelmann et al., 2019; Janneck et al., 2012). Da aufgrund der hierarchischen Struktur der Selbstkonzepte eine starke Korrelation zwischen computerbezogenem und physikbezogenem Selbstkonzept vermutet wurde, erfolgte im Hinblick auf mögliche Multikollinearität eine Pearson-Korrelationsanalyse. Es zeigte sich mit $r = .29$ unerwarteterweise eine mittlere Korrelation der Konstrukte (Cohen, 1988). Einen möglichen Erklärungsansatz für dieses Ergebnis stellt das Internal/External-Frame-of-Reference-Modell (I/E-Modell) dar, wonach Schüler:innen ihre akademischen Selbstkonzepte häufig an externen Faktoren wie Schulnoten festmachen bzw. diese Faktoren als Bezugsnorm verwenden (D. H. Rost et al., 2004). Fehlt es hier an Erfahrung bzw. Bezugsnormen, gestaltet sich eine Selbsteinschätzung als schwieriger. Im Gegensatz zu Physik, wo die Schule durch Feedback eine Bezugsnorm vorgibt, spielt sich der Computergebrauch häufig zu Hause ab und es fehlt der soziale Vergleich. Ein weiterer Grund könnte auch in der Formulierung der Items liegen. Beim computerbezogenen Selbstkonzept wurde vor allem die Erfahrung im Umgang mit Computern, der häufig zu Hause stattfindet, abgefragt. Demgegenüber lag bei den Items zum physikbezogenen Selbstkonzept der Fokus auf unterrichtsbezogenen Facetten. Während beim physikbezogenen Selbstkonzept für alle Schüler:innen eine gemeinsame Bezugsnorm vorhanden ist, existiert diese beim computerbezogenen Selbstkonzept nur für solche, die Erfahrung mit Computern haben, wovon nicht für beide Geschlechter gleichermaßen ausgegangen werden kann (Eickelmann et al., 2019; Vollmeyer & Imhof, 2007).

Eine hohe Korrelation ($r = .71$) konnte zwischen physikbezogenem Selbstkonzept und physikbezogener Selbstwirksamkeitserwartung ermittelt werden. Auch wenn der Wert unterhalb der Grenze für mögliche Multikollinearitätseffekte liegt, ist die Stärke des Zusammenhangs auffällig und könnte auf ein Problem der Skalen hindeuten, dass in den Limitationen erläutert werden soll.

6.2 Intervention

Zunächst konnte die erste Forschungsfrage positiv beantwortet werden, da das Serious Game in geringer Effektstärke ($d = .37$) mehr Flow-Erleben ermöglicht als die konventionellen Unterrichtsmaterialien der Kontrollgruppe. Die geringe Effektstärke ist ein Hinweis darauf, dass die Materialien der Kontrollgruppe ebenfalls die Möglichkeit bieten, Flow zu erleben. Dies kann bspw. darauf zurückgeführt werden, dass die Schüler:innen während der Intervention Möglichkeiten zum Durchführen von Versuchen und selbstgesteuerten Lernen hatten (Seidel et al., 2006). Dieses Erleben von Autonomie ist nicht nur ein Merkmal der Selbstbestimmungstheorie (Deci & Ryan, 1985), sondern steht, wie in Abschnitt 2.2 beschrieben, vermutlich in Zusammenhang mit dem Flow-Erleben (Tausch & Peifer, 2019). Bislang wurden nur wenige Studien mit einer adäquaten Kontrollgruppe durchgeführt (R. Mayer, 2019). Daher ist dieses Ergebnis entscheidend für die Forschung zu Serious Games, da es für die Vermutung spricht, dass Serious Games mit bestimmten Merkmalen, wie Flow-förderlichen Elementen und der Möglichkeit des selbstregulierten Lernens, auch Flow-Erleben auslösen.

Innerhalb der Experimental- und der Kontrollgruppe zeigten sich geschlechtsspezifische Unterschiede (a): In beiden Gruppen erlebten die Schüler mehr Flow als die Schülerinnen. Während dieser

Unterschied in der Kontrollgruppe mit mittlerer Effektstärke ($d = .78$) auftrat, war die Effektstärke in der Experimentalgruppe gering ($d = .38$). Dies kann als Hinweis interpretiert werden, dass die Gestaltung des Spiels, die nach den Empfehlungen aus der Literatur gezielt Frauen bzw. Mädchen ansprechen sollte (Alserri et al., 2018; Riopel et al., 2019), gelungen ist. Dennoch widerspricht dieses Ergebnis anderen Studien (Atwood-Blaine & Huffman, 2017; Bressler & Bodzin, 2016; Bressler et al., 2019; Inal & Cagiltay, 2007), in denen höhere Werte bei Schülerinnen als bei Schülern ermittelt wurden. Aus diesen Studien ging hervor, dass bei Schülerinnen und Schülern das Flow-Erleben durch jeweils andere Elemente gefördert wird. Beispielsweise bevorzugen Schüler kompetitive Elemente, während bei Schülerinnen die narrativen Elemente zum Flow-Erleben führen (Bressler & Bodzin, 2013, 2016). Im Serious Game der Intervention standen die Lerninhalte im Fokus, und die Narration beschränkte sich auf wenige Sequenzen, was dazu geführt haben könnte, dass Mädchen weniger Flow erlebten.

Die zweite Forschungsfrage sollte durch multiple lineare Regressionen beantwortet werden, um Erkenntnisse darüber zu gewinnen, welche der Variablen - *Intervention*, *Geschlecht*, *computerbezogenes* und *physikbezogenes Selbstkonzept*, *physikbezogene Selbstwirksamkeitserwartung(t_1)* sowie *Fachwissen(t_1)* - in einem signifikanten Zusammenhang mit dem Flow-Erleben stehen. Die Intensität des Flow-Erlebens wurde maßgeblich durch die Art der Intervention bestimmt ($B = -.25$, $p = .003$). Der Koeffizient weist darauf hin, dass der Effekt bei der Experimentalgruppe dabei stärker war. Dies deckt sich auch mit dem signifikant höheren Flow-Erleben in der Experimentalgruppe, das in der ersten Fragestellung ermittelt wurde. Auch zeigte sich in der linearen Regression der bereits beobachtete Trend, dass Schüler, dem Vorzeichen nach, allgemein mehr Flow erleben als Schülerinnen ($B = -.19$, $p = .037$). Weiterhin bestehen lineare Zusammenhänge zwischen dem physikbezogenen ($B = .21$, $p = .005$) und dem computerbezogenen ($B = .12$, $p = .018$) Selbstkonzept (b und c) mit dem Flow-Erleben. Dies ist nicht überraschend, da die prädiktive Eigenschaft des physikbezogenen und des computerbezogenen Selbstkonzepts für motivationale Zustände bereits ermittelt wurde (Janneck et al., 2012; D. H. Rost & Sparfeldt, 2002).

Überraschenderweise stellte sich heraus, dass die physikbezogene Selbstwirksamkeitserwartung(t_1) im multiplen Regressionsmodell nicht als signifikanter Prädiktor für das Erleben von Flow (e) fungiert. Auch in den Regressionsmodellen, die für die Experimentalgruppe und die Kontrollgruppe jeweils aufgestellt wurden, zeigte die Selbstwirksamkeitserwartung(t_1) keine signifikanten Prädiktoreigenschaften. Die bereits diskutierte Korrelation mit dem physikbezogenen Selbstkonzept, wodurch dieses als Suppressorvariable in der Regressionsgleichung erscheint (Bortz & Schuster, 2010), konnte jedoch aufgrund der Höhe der Korrelation ausgeschlossen werden. Darüber hinaus ist zu beachten, dass der Suppressoreffekt durch eine mögliche Multikollinearität nur die Signifikanz beeinflusst hätte, nicht jedoch den Koeffizienten. Dieser ist nahe null ($B < .01$), womit sein Beitrag zur gesamten Regressionsgleichung marginal ist. Es ist also davon auszugehen, dass die physikbezogene Selbstwirksamkeitserwartung in Zusammenhang mit den starken Prädiktoren *computerbezogenes* und *physikbezogenes Selbstkonzept* (b und c) keinen Einfluss auf das Flow-Erleben hat.

Die Prüfung des Regressionsmodells für die Experimentalgruppe zeigte, dass hier das Geschlecht in Kombination mit den anderen Variablen keinen signifikanten Einfluss auf das Flow-Erleben hat. Dies scheint den Ergebnissen von Welch's t-Tests aus der ersten Fragestellung zu widersprechen. Eine einfache lineare Regression mit dem Geschlecht als Prädiktor zeigte jedoch ebenfalls einen signifikanten Wert auf. Erst durch den linearen Zusammenhang mit anderen Prädiktoren verliert das Geschlecht seinen Einfluss und lediglich die Selbstkonzepte sagen das Flow-Erleben signifikant vorher. Daraus kann abgeleitet werden, dass das Flow-Erleben unabhängig vom Geschlecht ist und vermutlich

von der Einstellung gegenüber Computer(spiele)n und/oder dem Fach Physik sowie von den betreffenden Selbstkonzepten abhängt.

Anders gestaltet sich die Situation in der Kontrollgruppe: In dieser ist neben dem physikbezogenen Selbstkonzept ($B = -19$, $p = .049$) das Geschlecht ($B = -.34$, $p = .013$) ein signifikanter Prädiktor für das Flow-Erleben. Der Koeffizient kann so interpretiert werden, dass insbesondere den Schülern das Flow-Erleben möglich war. Dies bestätigt die Annahme, dass der Physikunterricht vor allem für Schüler interessant ist, während er Schülerinnen wenig anspricht (Strömmer & Winkelmann, 2020).

Nicht signifikant ist in beiden Gruppen der Einfluss des Fachwissens vor der Intervention (e) auf das Flow-Erleben. Diese Ergebnisse sind im Hinblick auf die Korrelationstabellen nicht überraschend, da auch hier kein signifikanter Zusammenhang mit dem Flow-Erleben ermittelt werden konnte. Bemerkenswert ist jedoch, dass sowohl die Ergebnisse der Korrelation als auch der Regressionstabelle negative Vorzeichen für das Fachwissen (t_1) in der Experimentalgruppe aufzeigen. Dies könnte wie folgt interpretiert werden: Je mehr Fachwissen bekannt war, desto weniger Flow erlebten Schüler:innen im Serious Game. Dies kann ein möglicher Hinweis sein, dass die in Kapitel 2.2 beschriebene *Passung aus Fähigkeiten und Herausforderung* (Csikszentmihalyi, 2014) in dem Serious Game entweder nicht vorhanden oder keine notwendige Bedingung für das Flow-Erleben war. Zu ähnlichen Ergebnissen kamen auch andere Studien (Rachmatullah et al., 2021). Eine weitere Erklärung ist, dass die Passung nicht mit den Inhalten des Serious Games in Zusammenhang stehen muss, sondern mit etwas anderem wie den Fähigkeiten im Umgang mit dem Computer.

6.3 Posttest

Die dritte Fragestellung zur Wirkung des Flow-Erlebens auf die Veränderung von Fachwissen und Selbstwirksamkeitserwartung wurde mit Strukturgleichungs-, Moderations- und Mediationsmodellen untersucht. Hinsichtlich des Fachwissens (t_2) profitierten sowohl die Experimentalgruppe (mit schwacher Effektstärke $d = .36$) als auch die Kontrollgruppe (mit mittlerer Effektstärke $d = .51$), wobei keine signifikanten Unterschiede zwischen den Geschlechtern festgestellt wurden. Dies stimmt mit den Ergebnissen anderer Studien überein, in denen ebenfalls keine höhere Lernwirksamkeit durch Serious Games ermittelt werden konnte (Furió et al., 2013; Wouters et al., 2013). Das ist ein Hinweis darauf, dass es möglicherweise andere Mechanismen gibt, die die Lernwirksamkeit beeinflussen, wie wissenschaftliche Intelligenz oder Interesse (Rosenthal et al., 2022). Beispielsweise konnten Sitzmann und Ely (2011) in ihrer Metaanalyse zeigen, dass Kontrollgruppen mit aktivem Lernmaterial mehr lernen ($d = .19$) als Experimentalgruppen mit Serious Games, woraus zu schließen ist, dass das Aktivitätsniveau entscheidend ist und nicht die Unterrichtsmethode (Riopel et al., 2019).

Auch andere Erklärungsansätze sind hier möglich, bspw. dass eine zu hohe Immersion in der Experimentalgruppe mit einer geringen Konzentration einherging (Böhme et al., 2020; Ke & Abras, 2013; Ravysse et al., 2017; Virvou et al., 2005).

Schließlich muss auch der Fachwissenstest selbst in Betracht gezogen werden. Er zeigte sich als insgesamt schwierig, vor allem für die Schüler:innen der Realschulen. Des Weiteren war die Prüfung der Reliabilität nicht zufriedenstellend durchführbar, sodass auch fraglich ist, ob die Reihenfolge von A- und B- Test die Ergebnisse beeinflusste. Um ein zuverlässiges Ergebnis zu erreichen, sollte der Fachwissenstest einer erneuten Prüfung unterzogen, angepasst und anschließend erneut eingesetzt werden oder ein neuer Test konzipiert werden.

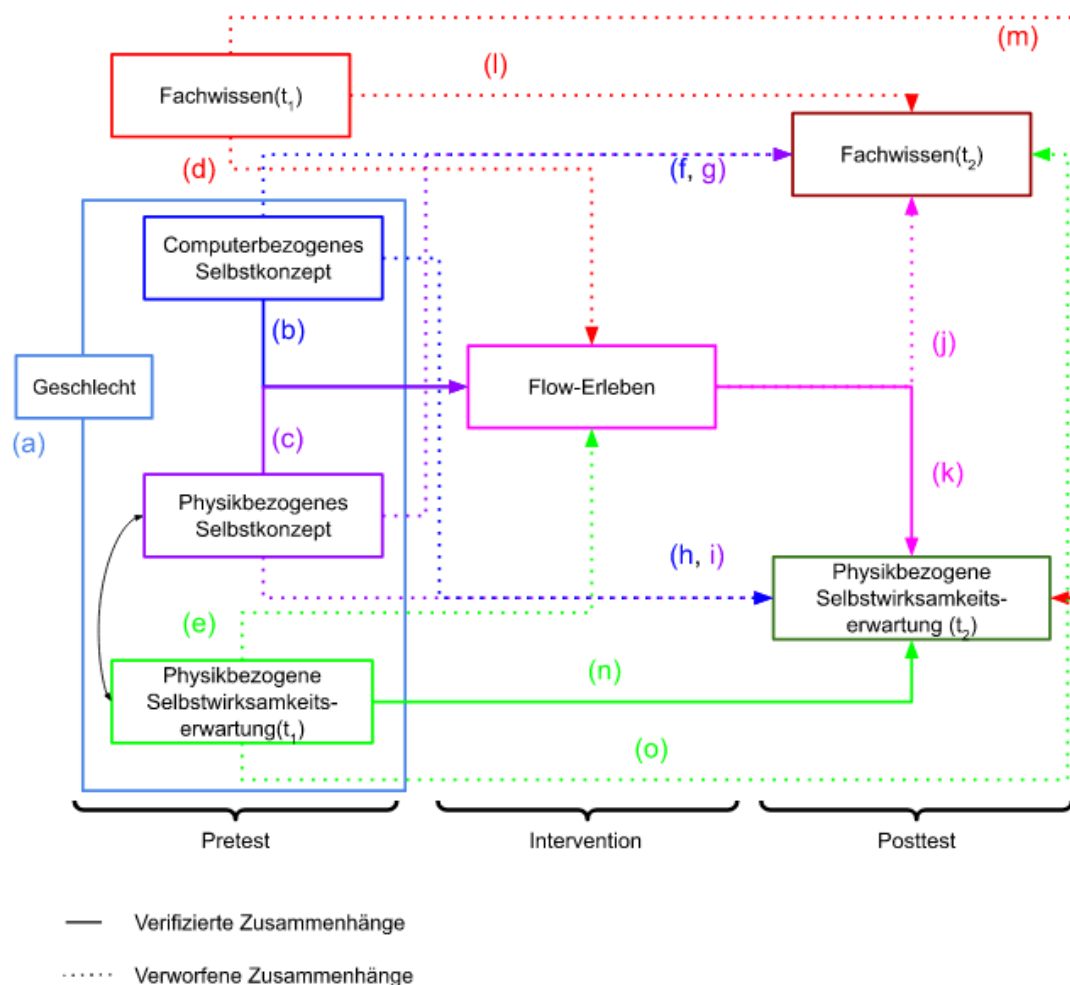
Vielversprechend sind die Ergebnisse zur physikbezogenen Selbstwirksamkeitserwartung (t_2) nach der Intervention. Sie bestätigen die Forschung von Ketelhut (2007) und (Sitzmann, 2011), wonach sich die Selbstwirksamkeitserwartung durch Serious Games verbessert. So profitierte die Experimentalgruppe

signifikant von der Intervention, wenn auch mit schwacher Effektstärke ($d = .25$, Schüler $d = .27$, Schülerinnen $d = .24$). Die Kontrollgruppe dagegen profitierte hinsichtlich der Selbstwirksamkeitserwartung nicht. Das Serious Game hat möglicherweise das Potential, die physikbezogene Selbstwirksamkeitserwartung zu stärken.

6.3.1 Zusammenhänge im Strukturgleichungsmodell der Experimentalgruppe

Inwiefern die Variablen des Pretests sowie das Flow-Erleben während der Intervention dazu beitragen, dass Schüler:innen lernen und sich als selbstwirksam erleben, wurde in einem Strukturgleichungsmodell untersucht. Die Ergebnisse für die Experimentalgruppe sollen anhand der Abbildung 25 diskutiert werden.

Abbildung 25 Nomologisches Netzwerk für die Ergebnisse der Experimentalgruppe.



Das Geschlecht (a) zeigte seinen Einfluss auf das physikbezogene ($B = -.36, p = .002, 95\%-CI [-.58, -.14]$) und computerbezogene ($B = -.80, p < .001, 95\%-CI [-1.00, -.50]$) Selbstkonzept und die physikbezogenen Selbstwirksamkeitserwartung ($B = -.29, p = .025, 95\%-CI [-.55, -.03]$). Die Wirkung des computerbezogenen (b, $B = .12, p = .045, 95\%-CI [.00, .27]$) und des physikbezogenen Selbstkonzepts (c, $B = .28, p = .014, 95\%-CI [.06, .49]$) zum Flow-Erleben war in diesem Modell erwartungsgemäß signifikant. Zwischen den Selbstkonzepten (f und g) und dem Fachwissen(t₂) konnte kein signifikanter Zusammenhang ermittelt werden. Keine der Theorien zur Veränderung des Fachwissens im Bereich der Elektrizitätslehre durch die Intervention mit einem Serious Game erwies sich als zutreffend. Auch

eine Überprüfung auf eine mögliche Moderatorfunktion ergab kein Modell, das zu den Daten passte: Der Interaktionsterm aus der Moderationsanalyse von Flow-Erleben und Fachwissen (t_1) wurde nicht signifikant. Auch ein Mediationseffekt des Flow-Erlebens wie bei Hoblitz (2015) konnte nicht ermittelt werden. Dies lässt sich, wie bereits erläutert, auf die schlechte Reliabilität und die hohe Schwierigkeit des Fachwissenstests zurückführen. Aber es kann auch ein weiterer Hinweis dafür sein, dass es andere Mechanismen der Lernwirksamkeit gibt, die in diesem Zusammenhang nicht erfasst wurden. Wie J. Wang et al. (2022) herausfanden, ist das Fachwissen vor der Intervention für die Lernwirksamkeit des Spiels nicht von Bedeutung. Aber auch nicht erhobene Konstrukte wie naturwissenschaftliches oder technisches Interesse könnten hier die Zusammenhänge erklären (Rosenthal et al., 2022). Schließlich besteht eine weitere mögliche Erklärung darin, dass sich Schüler:innen mit hohem Fachwissen(t_1) im Fachwissenstest nach der Erhebung nicht mehr verbessern konnten, was auch in anderen Studien bereits herausgefunden wurde (Rosenthal et al., 2022)(ebd.). Dies bedeutet unter Umständen, dass der Zusammenhang nicht linear ist, sodass die lineare Regression nicht die passende Methode ist. Dies wird in den Limitationen im folgenden Abschnitt aufgegriffen und es sollen mögliche Alternativen aufgezeigt werden.

Signifikante Zusammenhänge konnten bezüglich der Selbstwirksamkeitserwartung(t_2) aufgedeckt werden. Zum einen zeigte sich ein starker Zusammenhang zwischen der Selbstwirksamkeitserwartung(t_1) und Selbstwirksamkeitserwartung(t_2) (n , $B = .52$, $p < .001$, $95\%-CI [.38, .67]$), aber auch das Flow-Erleben (k , $B = .27$, $p < .001$, $95\%-CI [.19, .42]$) beeinflusst die Selbstwirksamkeitserwartung(t_2). Eine Mediationsanalyse zeigte keinen signifikanten Interaktionsterm zwischen der Selbstwirksamkeitserwartung(t_1) und Selbstwirksamkeitserwartung(t_2) und dem Flow-Erleben, jedoch erwies sich die multiple Regression der Variablen als signifikant und konnte mit 84 % einen hohen Anteil (Cohen, 2013) der Varianz erklären. Auch in der Kontrollgruppe zeigte sich ein ähnliches Ergebnis. Die Annahmen eines Mediationseffekts (Pavlas et al., 2010) konnten weder für die Experimentalgruppe noch für die Kontrollgruppe bestätigt werden. Somit besteht die Möglichkeit, dass hohes Flow-Erleben und hohe Selbstwirksamkeitserwartung gemeinsam auftreten, da sie durch ähnliche Komponenten der Settings hervorgerufen werden. Darüber hinaus bestätigt sich dadurch die Annahme, dass eine hohe Selbstwirksamkeitserwartung mit hohem Flow-Erleben einhergeht und dieses wiederum mit einer hohen Selbstwirksamkeitserwartung (Rodríguez-Sánchez et al., 2011; Salanova et al., 2006; Zubair & Kamal, 2015). Dabei zeigten die nicht signifikanten Ergebnisse aus der Moderationsanalyse auch, dass diese beiden Konstrukte nicht in Interaktion stehen. Ein bemerkenswertes Ergebnis ist, dass das Geschlecht keinen Einfluss auf das Fachwissen(t_2) sowie die Selbstwirksamkeitserwartung(t_2) zeigte, woraus geschlossen werden kann, dass das Serious Game durch seine Gestaltung gemäß den Hinweisen aus der Literatur (Alserrri et al., 2018) die Unterschiede aus dem Pretest ausgleichen konnte.

6.3.2 Zusammenhänge im Strukturgleichungsmodell der Kontrollgruppe

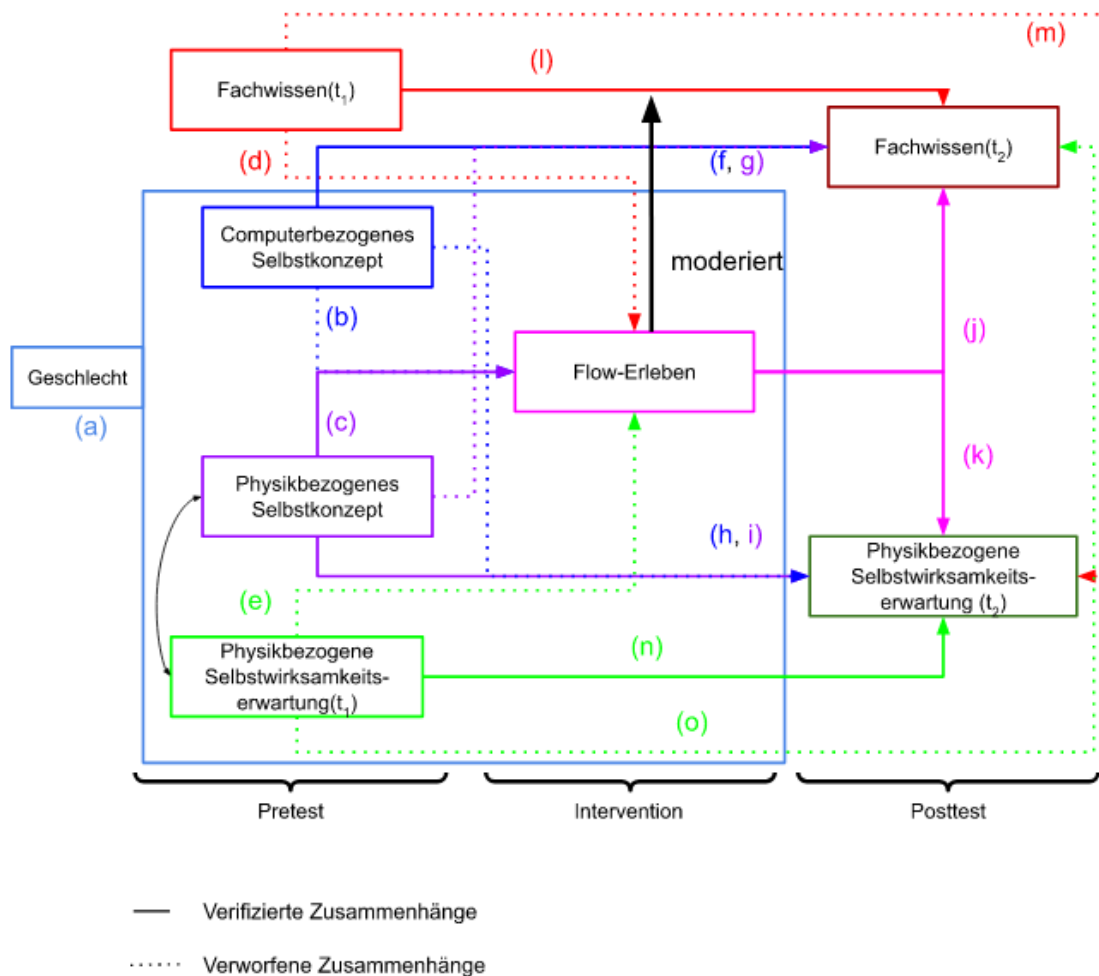
Das gleiche Strukturgleichungsmodell zeigte mit den Daten der Kontrollgruppe einige Unterschiede zur Experimentalgruppe. Der signifikante Einfluss des Geschlechts über die selbstbezogenen Einstellungen hinaus auch auf das Flow-Erleben (a , $B = .33$, $p = .012$, $95\%-CI [.59, .07]$) konnte gezeigt werden. Das Geschlecht zeigte keinen Einfluss auf das Fachwissen(t_2) sowie die Selbstwirksamkeitserwartung(t_2), was möglicherweise auf die Gestaltung der Materialien der Kontrollgruppe zurückzuführen ist, die allen Schüler:innen gleichermaßen ein Experimentieren in einer Umgebung ohne Bewertung und Druck ermöglichte. Ein unerwartetes Ergebnis ist die signifikante Prädiktoreigenschaft des computerbezogenen Selbstkonzepts (f , $B = -.04$, $p = .032$, $95\%-$

CI [-.07, .00]) für das Fachwissen(t_2), insbesondere das negative Vorzeichen wirft dabei Fragen auf. So scheint es, dass vor allem Schüler:innen mit geringem computerbezogenen Selbstkonzept bezüglich des Fachwissens profitieren. Dies könnte ein Hinweis sein, dass Schüler:innen, die sich im Umgang mit Computern unsicher fühlen, von traditionellem Physikunterricht stärker profitieren.

Das bereits vorhandene Fachwissen(t_1) zeigte eine signifikante Wirkung auf das Fachwissen(t_2) in der Kontrollgruppe(l, $B = -.50$, $p < .001$, 95%-CI [.30, .69]). Dies im Zusammenhang mit dem ebenfalls signifikanten Effekt des Flow-Erlebens auf das Fachwissen(t_2) (j, $B = .04$, $p = .034$, 95%-CI [.00, .09]) wurde in einer Moderationsanalyse überprüft. Insbesondere wurde der Interaktionsterm aus Fachwissen (t_1) und Flow-Erleben signifikant ($B = .29$, $p = .043$, 95%-CI [.01, .57]). Dies trat bereits bei unterdurchschnittlichen Werten für das Flow-Erleben auf. So lag der Mittelwert der erhobenen Werte des Flow-Erlebens in der Kontrollgruppe bei $M = 3.15$ ($SD = .71$), dem Jonson-Neyman-Plot in Abbildung 24 konnte aber entnommen werden, dass schon ab einem Flow-Erleben mit Zustimmungswerten von $X = 2.41$ der Effekt von Fachwissen(t_1) auf Fachwissen(t_2) signifikant wurde. Das Flow-Erleben verstärkt folglich die Wirkung des Vorwissens auf die Steigerung des Fachwissens, was zu der Vermutung führt, dass die Materialien gerade für Schüler:innen mit Vorwissen motivierend und lernförderlich waren. Eine mögliche Schlussfolgerung besteht darin, dass durch solche Unterrichtsszenarien vor allem diejenigen Schüler:innen gefördert werden, die ohnehin Freude am Physikunterricht haben und entsprechend profitieren.

Die Selbstwirksamkeitserwartung(t_2) wurde neben dem Flow-Erleben (k, $B = .25$, $p < .001$, 95%-CI [.13, .38]) und der Selbstwirksamkeitserwartung(t_1) (n, $B = .50$, $p < .001$, 95%-CI [.36, .63]) noch vom physikbezogenen Selbstkonzept (i, $B = .23$, $p = .001$, 95%-CI [.09, .37]) vorhergesagt. Der Zusammenhang zwischen physikbezogenem Selbstkonzept und physikbezogener Selbstwirksamkeitserwartung ist nicht verwunderlich, da Schüler:innen in der Kontrollgruppe mit einer positiven Einstellung die Aufgaben bearbeitet haben und bei eventuellen Misserfolgen frustationstoleranter waren (R. Schwarzer & Jerusalem, 2002). So konnten sie sich als wirksam erleben und ihre Selbstwirksamkeitserwartung steigern.

Abbildung 26 Nomologisches Netzwerk für die Ergebnisse der Kontrollgruppe.



6.4 Limitation und Ausblick

Zum Schluss soll auf die in der Diskussion angedeuteten Limitationen der Arbeit eingegangen und ein Ausblick gegeben werden. Dabei sollen zunächst mögliche Limitationen durch die Materialien und anschließend durch die Methoden erläutert werden.

Bezüglich der Testinstrumente existieren in der Theorie Unklarheiten hinsichtlich der tatsächlichen Rolle der verschiedenen Komponenten des Flow-Erlebens (Abuhamdeh, 2020). Ein Problem besteht darin, dass in vielen Arbeiten das Flow-Erleben immer noch ausschließlich über die erste Komponente *Passung von Herausforderung und Fähigkeiten* operationalisiert wird (Csikszentmihalyi & LeFevre, 1989; Engeser & Rheinberg, 2008; Massimini & Carli, 1988; Shernoff et al., 2003). Es herrscht jedoch Uneinigkeit darüber, wie die Gewichtung der Passung sein sollte. In einigen Studien wird angenommen, dass ein Gleichgewicht zwischen Fähigkeit und Herausforderung das Flow-Erleben verursacht (C.-C. Wang & Hsu, 2014), während in anderen eine etwas höhere Herausforderung als entscheidend für das Flow-Erleben betrachtet wird (Abuhamdeh & Csikszentmihalyi, 2012; Barthelmäs & Keller, 2021; Moneta & Csikszentmihalyi, 1999). Es besteht auch Uneinigkeit darüber, ob das Flow-Erleben als diskretes Konstrukt betrachtet werden sollte (d. h., ob jemand Flow erlebt oder nicht), oder ob es sich um ein kontinuierliches Konstrukt (es wird mehr oder weniger Flow während einer Handlung erlebt) handelt (Abuhamdeh, 2020). In der Mehrheit der Forschungsarbeiten

zu Flow wird es als kontinuierliches Konstrukt aufgefasst (ebd.), was dazu führt, dass die Zeitpunkte zu denen Flow erhoben wird, in zukünftigen Modellen beachtet werden sollte.

Eine weitere Limitation ist die starke Korrelation zwischen dem physikbezogenen Selbstkonzept und der physikbezogenen Selbstwirksamkeitserwartung. In ihrer Studie verwenden Marsh et al. (2019) den Begriff *Jingle-Jangle-Trugschluss* (Kelley 1927) für das Phänomen, dass entweder über Skalen mit derselben Bezeichnung unterschiedliche Konstrukte gemessen werden (Jingle) oder mit zwei Skalen gleiche Konstrukte (Jangle). Da diese Konstrukte bei dieser spezifischen Ausprägung ähnliche Aspekte messen (Marsh et al., 2019), sollte in zukünftigen Untersuchungen darauf geachtet werden, dass die Skala zur Messung des Selbstkonzepts und jene für die Selbstwirksamkeitserwartung unterschiedlich sind. Eine Möglichkeit ist z. B. die Untersuchung mittels konfirmatorischer Faktorenanalyse, bei der die zugrunde liegende Struktur mit den zwei Faktoren *Selbstkonzept* und *Selbstwirksamkeitserwartung* überprüft wird (ebd.). Da der Jangle-Trugschluss nicht ausgeschlossen werden kann, müssen entsprechende Skalen in zukünftigen Arbeiten zuerst einer Faktorenanalyse unterzogen werden. Sollten sich dabei die Skalen zu *Selbstwirksamkeitserwartung* und *Selbstkonzept* nicht als getrennte Faktoren erweisen, ist davon auszugehen, dass die beiden Skalen ein gemeinsames Konstrukt abbilden (Bong & Skaalvik, 2003).

Ein weiteres Problem betrifft die hohe Schwierigkeit des Fachwissenstests. Die Wahrscheinlichkeit, dass die Items gelöst wurden, war mittelhoch. Hinsichtlich der Schwierigkeit wird empfohlen, dass die Lösungswahrscheinlichkeit der Items zwischen .2 und .8 liegt (Bühner, 2011), während in dieser Arbeit Werte von .1 bis .6 erreicht wurden. Die Schwierigkeit der Elektrizitätslehre liegt unter anderem darin, dass Schüler:innen zum ersten Mal mit den damit zusammenhängenden Begriffen konfrontiert werden, was eine hohe kognitive Belastung darstellt. Sie müssen nicht nur Fachbegriffe lernen, sondern auch die dahinterstehenden Konzepte verstehen. Außerdem ist es erforderlich, Alltagsbegriffe neu zu definieren. Es wurde bereits im Vorfeld darauf geachtet, die Formulierungen des Testinstruments anzupassen, indem z. B. Begriff wie *Lämpchen* oder *Birnchen* durch *Glühlampe* ersetzt wurden und angleichen der Schaltskizzen, sodass z. B. die Batterie immer die gleiche Polung aufwies. Problematisch beim Fachwissenstest ist darüber hinaus, dass die in früheren Arbeiten angegebene Güte des Testinstruments nicht nachvollzogen werden konnte. In früheren Arbeiten mit diesem Instrument wurde die Reliabilität mittels Cronbachs α berechnet wurde, wobei dazu alle 26 ursprünglichen Items verwendet wurden. Diese Reliabilitätsprüfung ist für das Messinstrument nicht geeignet, da bei Cronbachs α von einer eindimensionalen Skala ausgegangen wird. Das trifft bei einem Fachwissenstest, mit dem verschiedene Wissensbereiche und Fehlvorstellungen abgefragt werden, nicht zu. Darüber hinaus neigt Cronbachs α bei einer großen Anzahl von Items wie in der vorliegenden Arbeit dazu, die Reliabilität zu überschätzen. Eine weitere Arbeit mit dem Fragebogen nutzte dessen Strukturierung nach Schüler:innenvorstellungen als Grundlage für eine Reliabilitätsschätzung und erhielt ebenfalls gute Ergebnisse für die Reliabilität (Urban-Woldron & Hopf, 2012). Eine legitime Prüfung der Reliabilität ist die Retest-Reliabilität, bei der kontrolliert wird, ob ein Fachwissenstest zu zwei Zeitpunkten ähnliche Ergebnisse liefert. In der vorliegenden Arbeit konnte diese Überprüfung nicht mit ausreichender statistischer Power durchgeführt werden, wodurch die Interpretation der Ergebnisse zum Fachwissenstest kritisch betrachtet werden müssen. Daher sollte er in einer ausreichend großen Stichprobe erneut überprüft werden. In der Auswertung zeigte sich zudem, dass das Fachwissen (t_2) in der Experimentalgruppe durch keinen Parameter signifikant vorhergesagt wurde. Weder Moderationsanalysen noch Mediationsanalysen konnten weitere Erkenntnisse für die Experimentalgruppe liefern. Eine weitere Möglichkeit der Auswertung sind bspw. nicht-lineare

Verfahren, wie Ridge- oder Lasso-Regressionen (Hastie et al., 2001), die in Zukunft weiter untersucht werden sollten.

Eine weitere Limitation, die mit einem Ausblick verbunden ist, betrifft die Modellierung des Pre-Post-Modells dieser Untersuchung. Eine Alternative ist die Modellierung mithilfe des Gain-Score-Ansatzes. Dabei wird die Differenz zwischen Fachwissen(t_2) und Fachwissen(t_1) bzw. Selbstwirksamkeitserwartung(t_2) und Selbstwirksamkeitserwartung(t_1) gebildet sowie die Wirkung der Variablen auf diese Differenz ermittelt. Ein Vorteil dieses Ansatzes ist seine Einfachheit, allerdings wird insbesondere bei geringer Reliabilität des Messinstruments seine Zuverlässigkeit bezweifelt (Gliner et al., 2003). Auch eine Kovarianzanalyse ist ein möglicher Ansatz bei Pre-Post-Designs mit Kontrollgruppe. Dabei werden die Unterschiede in den Pretest-Ergebnissen durch Anpassungen der Fehlervarianz im Posttest berücksichtigt. Der Pretest fließt dabei als Kovariate ein. Dieser Test setzt neben der Homogenität der Regressionssteigung auch eine signifikante lineare Beziehung zwischen dem Pretest und den Posttest-Variablen voraus (Gliner et al., 2003). Das war, zumindest in dieser Untersuchung, für die Variable Fachwissen nicht gegeben.

Auch das verwendete Strukturgleichungsmodell muss kritisch betrachtet werden, da hier ausschließlich manifeste Variablen verwendet wurden, wurde die Streuung bzw. Varianz, die durch die Erhebung der einzelnen Items jedes Konstrukt nicht berücksichtigt. Jedoch wird für ein Modell mit latenten Variablen eine größere Stichprobe gefordert, worauf bei zukünftiger Forschung geachtet werden muss. Schließlich muss auch die Durchführung der Intervention selbst kritisch betrachtet werden, die zum einen nur an Schulen in Baden-Württemberg stattfand und zusätzlich durch die Pandemie erschwerte Suche nach geeigneten Schulen. Um die statistische Power solcher Untersuchungsdesigns zu verbessern, wäre eine größere Stichprobe vorteilhaft. Aber auch die unterschiedlichen Bedingungen, unter denen die Schüler:innen die Zeit des Fernunterrichts erlebten. Während für manche Schüler:innen diese Zeit eine gute Möglichkeit war, ihre selbstregulierten Lernfähigkeiten weiterzuentwickeln, stellte sie für andere Kinder eine Herausforderung dar. Interventionen, bei denen sie erneut eigenständig lernen mussten, wurden möglicherweise negativ bewertet. Es ist auch noch unklar, ob Proband:innen mit verschiedenen Selbstkonzepten unterschiedlich mit dem Serious Game interagiert haben. In diesem Zusammenhang ist eine detaillierte Analyse der Prozessdaten erforderlich.

7 Fazit

Die vorliegende Studie konnte in einem Pre-Post-Design mit Kontrollgruppe Hinweise für die Wirkungen eines Serious Games im Bereich Elektrizitätslehre auf das Flow-Erleben, Lernen und die Selbstwirksamkeitserwartung aufdecken. Insgesamt konnten die Daten von 310 Schüler:innen in der Auswertung berücksichtigt werden, von denen 168 der Experimentalgruppe zugeordnet wurden. Zwischen den 158 männlichen und 152 weiblichen Schüler:innen zeigten sich Unterschiede bei den im Pretest erhobenen Variablen. So zeigten sich höhere Werte bei physikbezogenem und computerbezogenem Selbstkonzept, sowie der physikbezogenen Selbstwirksamkeitserwartung. Keine Unterschiede dagegen gab es zwischen den Geschlechtern bezüglich des vorhandenen Fachwissens. Durch Regressionsmodelle konnte ermittelt werden, dass das computerbezogene und das physikbezogene Selbstkonzept wichtige Prädiktoren für das Flow-Erleben beim Spielen des Serious Games sind. Schüler:innen mit positiven Selbstkonzepten waren in der Lage, viel Flow zu erleben. Dies wurde über die Eigenschaft des Selbstkonzepts, Motivation vorherzusagen begründet. Serious Games, die unterschiedliche Selbstkonzepte berücksichtigen, können möglicherweise Flow-Erleben ermöglichen. In der Kontrollgruppe zeigte sich das physikbezogene Selbstkonzept im Rahmen einer Strukturgleichungsmodellierung ebenfalls als Prädiktor für das Flow-Erleben, während das computerbezogene Selbstkonzept hierbei keinen Zusammenhang aufwies. Keinen Einfluss auf das Flow-Erleben hatte in beiden Gruppen die physikbezogene Selbstwirksamkeitserwartung. Ein wichtiger Unterschied zwischen den beiden Gruppen betraf jedoch das Geschlecht. Obgleich Schüler höhere Werte beim Flow-Erleben angaben als Schülerinnen, zeigte das Geschlecht im Regressionsmodell der Experimentalgruppe keinen Einfluss auf das Flow-Erleben. Dagegen wurde das Flow-Erleben im Regressionsmodell der Kontrollgruppe vom Geschlecht - zu Gunsten der Schüler - beeinflusst. Der erlebte Flow war in beiden Gruppen ein signifikanter Prädiktor für die Höhe der Selbstwirksamkeitserwartung, die nach der Intervention gemessen wurde. Nur in der Kontrollgruppe konnte das Flow-Erleben darüber hinaus noch das Fachwissen vorhersagen und moderierte dessen Veränderung. In der Experimentalgruppe konnte durch keine der erhobenen Variablen das Fachwissen nach der Intervention vorhergesagt werden, sodass hier noch Bedarf für weitere Forschung ist.

Zusammenfassend konnten die Ergebnisse der Arbeit Hinweise darauf geben, wie die Selbstwirksamkeitserwartung durch geeignete Unterrichtsszenarien verbessert werden kann. Merkmale solcher Unterrichtsszenarien sind, dass sie einerseits die Theorie der Selbstwirksamkeitserwartung (Bandura, 1977) und andererseits Elemente des Flow-Erlebens berücksichtigen (Csikszentmihalyi, 2014). Solche Unterrichtsszenarien sind jedoch im Schulalltag schwer umsetzbar, da sie einen hohen Zeit- und Materialaufwand erfordern. Darüber hinaus fehlt es den Lehrkräften häufig an Möglichkeiten, den Unterricht auf die einzelnen Schüler:innen und ihre Fähigkeiten abzustimmen, um das Flow-Erleben zu ermöglichen. Serious Games bieten die Möglichkeit Flow-Erleben auszulösen und die Selbstwirksamkeitserwartung zu verbessern. Zudem sind sie ebenso effektiv für das Lernen wie konventionelle Unterrichtsszenarien, die den Schüler:innen die Möglichkeit des selbstregulierten Lernens bieten. Ein weiterer Vorteil von Serious Games besteht darin, dass sie im Vergleich mit anderen Unterrichtsmaterialien günstiger und zeiteffizient sind, da keine Experimente aufgebaut sowie differenzierte Übungen erstellt, kopiert und verteilt werden müssen. Jeder Schüler und jede Schülerin erhält genau die Hilfe, die er oder sie benötigt, und bekommt bei Bedarf direktes Feedback. Dies erhöht auch die Effizienz des Unterrichts, insbesondere wenn die Lehrkraft nicht die Möglichkeit oder Fähigkeit hat, alle Schüler:innen gleichermaßen durch verbale

Ermutigung zu erreichen. Serious Games erleichtern für die Lehrkraft die Auswertung von Ergebnissen und Prozessdaten. Zudem ermöglichen sie eine genauere Überwachung des Lernfortschritts und eine bessere Berücksichtigung der Vielfalt, um individuelle Unterstützung anzubieten. Die Bedeutung dieser Studienergebnisse liegt in den aufgezeigten Indizien, dass eine gesteigerte Selbstwirksamkeitserwartung Frauen potenziell dazu inspirieren könnte, berufliche Wege in MINT-Fachgebieten einzuschlagen. Durch die Erfahrung, selbstwirksam zu sein, haben Frauen mehr Vertrauen in ihre Fähigkeit, schwierigere Aufgaben im MINT-Bereich zu bewältigen, was wiederum zu höheren Erwartungen an die Selbstwirksamkeit führt. Durch die Schaffung einer positiven Lernumgebung, in der die Interessen und Bedürfnisse der Mädchen berücksichtigt werden, kann ein wesentlicher Beitrag zur Förderung von Mädchen im MINT-Bereich geleistet werden.

Literaturverzeichnis

- Abt, C. C. (1971). *Ernstes Spiel: lernen durch gespielte Wirklichkeit* (1. Aufl.). Kiepenheuer & Witsch.
- Abuhamdeh, S. (2020). Investigating the "Flow" Experience: Key Conceptual and Operational Issues. *Frontiers in psychology, 11*, Artikel 158. <https://doi.org/10.3389/fpsyg.2020.00158>
- Abuhamdeh, S. (2021). On the Relationship Between Flow and Enjoyment. In C. Peifer & S. Engeser (Hrsg.), *Advances in Flow Research* (S. 155–169). Springer; Springer International Publishing. https://doi.org/10.1007/978-3-030-53468-4_6
- Abuhamdeh, S. & Csikszentmihalyi, M. (2012). The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities. *Personality & social psychology bulletin, 38*(3), 317–330. <https://doi.org/10.1177/0146167211427147>
- Admiraal, W., Huizenga, J., Akkerman, S. & Dam, G. t. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194. <https://doi.org/10.1016/j.chb.2010.12.013>
- Alexiou, A. & Schippers, M. C. (2018). Digital game elements, user experience and learning: A conceptual framework. *Education and Information Technologies, 23*(6), 2545–2567. <https://doi.org/10.1007/s10639-018-9730-6>
- Alserri, S. A., Zin, N. A. M. & Wook, T. S. M. T. (2018). Gender-based Game Engagement Model Validation using Low Fidelity Prototype. *International Journal of Engineering and Advanced Technology, 8*(4), Artikel D9042049420, 1350–1357. <https://doi.org/10.35940/ijeat.D9042.049420>
- Annetta, L. A., Frazier, W. M., Folta, E., Holmes, S. Y., Lamb, R. L. & Cheng, M.-T. (2013). Science Teacher Efficacy and Extrinsic Factors Toward Professional Development Using Video Games in a Design-Based Research Model: The Next Generation of STEM Learning. *Journal of Science Education and Technology, 22*(1), 47–61. <https://doi.org/10.1007/s10956-012-9375-y>
- Anscombe, F. J. & Glynn, W. J. (1983). Distribution of the kurtosis statistic b_2 for normal samples. *Biometrika, 70*(1), 227–234. <https://doi.org/10.1093/biomet/70.1.227>
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B. & Wong, B. (2012). "Balancing acts": Elementary school girls' negotiations of femininity, achievement, and science. *Science Education, 96*(6), 967–989. <https://doi.org/10.1002/sce.21031>
- Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F [Francesco], Freitas, S. de, Louchart, S., Suttie, N., Berta, R [Riccardo] & Gloria, A. de (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology, 46*(2), 391–411. <https://doi.org/10.1111/bjet.12113>
- Atwood-Blaine, D. & Huffman, D. (2017). Mobile gaming and student interactions in a science center: the future of gaming in science education. *International journal of science and mathematics education, 15*, 45–65.
- Backhaus, K., Erichson, B., Gensler, S., Weiber, R. & Weiber, T. (2021). *Multivariate Analysemethoden*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-32425-4>
- Bai, S., Hew, K. F. & Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review, 30*, Artikel 100322. <https://doi.org/10.1016/j.edurev.2020.100322>
- Baltes-Götz, B. (1994). Einführung in die Analyse von Strukturgleichungsmodellen mit LISREL 7 und PRELIS unter SPSS. *Universität Trier: Universitäts-Rechenzentrum, Trier*.

- Bandalos, D. L. & Finney, S. J. (2002). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Hrsg.), *Advanced structural equation modeling: (269–296)*. Lawrence Erlbaum.
- Bandalos, D. L., Yates, K. & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology, 87*(4), 611–623. <https://doi.org/10.1037/0022-0663.87.4.611>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.
- Bandura, A. & Cervone, D. (1986). Differential engagement of self-reactive influences in cognitive motivation. *Organizational behavior and human decision processes, 38*(1), 92–113. [https://doi.org/10.1016/0749-5978\(86\)90028-2](https://doi.org/10.1016/0749-5978(86)90028-2)
- Bandura, A. & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of applied psychology, 88*(1), 87.
- Bandura, A. & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of personality and social psychology, 41*(3), 586–598. <https://doi.org/10.1037/0022-3514.41.3.586>
- Barthelmäs, M. & Keller, J. (2021). Antecedents, Boundary Conditions and Consequences of Flow. In C. Peifer & S. Engeser (Hrsg.), *Advances in Flow Research* (S. 71–107). Springer; Springer International Publishing. https://doi.org/10.1007/978-3-030-53468-4_3
- Bartosch, I. (2013). *Entwicklung weiblicher Geschlechtsidentität und Lernen von Physik - ein Widerspruch? // Entwicklung weiblicher Geschlechtsidentität und Lernen von Physik - ein Widerspruch?* Teilw. zugl.: 2011, Klagenfurt, Univ., Diss., 2011 (Online-Ausg). *Internationale Hochschulschriften: Bd. 598*. Waxmann Verlag.
- Barzel, B., Reinhoffer, B. & Schrenk, M. (2012). Das Experimentieren im Unterricht. *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht. Schüler lernen wissenschaftlich denken und arbeiten*, 103–127.
- Barzilai, S. & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education, 70*, 65–79. <https://doi.org/10.1016/j.compedu.2013.08.003>
- Beierlein, C. (2011). *Geschlechtsunterschiede in Motivation und Strategiesystematik beim selbstregulierten Erlernen eines Computerprogramms* [Dissertation]. Goethe-Universität Frankfurt am Main, Frankfurt am Main.
- Bellotti, F [F.], Berta, R [R.], Gloria, A. de, Lavagnino, E., Dagnino, F., Ott, M., Romero, M., Usart, M. & Mayer, I. (2012). Designing a Course for Stimulating Entrepreneurship in Higher Education through Serious Games. *Procedia Computer Science, 15*, 174–186. <https://doi.org/10.1016/j.procs.2012.10.069>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin, 107*(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bentler, P. M. & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78–117.
- Berger, M. (2018). *Neue Medien im experimentellen Physikunterricht der Sekundarstufe I - eine empirisch-explorative Studie zur Untersuchung der Auswirkungen von virtuellem Lernen* [Dissertation]. Pädagogische Hochschule Heidelberg.
- Bergey, B. W., Ketelhut, D. J., Liang, S., Natarajan, U. & Karakus, M. (2015). Scientific Inquiry Self-Efficacy and Computer Game Self-Efficacy as Predictors and Outcomes of Middle School Boys' and Girls' Performance in a Science Assessment in a Virtual Environment. *Journal of Science Education and Technology, 24*(5), 696–708. <https://doi.org/10.1007/s10956-015-9558-4>

- Blumberg, F. C. & Fisch, S. M. (2013). Introduction: digital games as a context for cognitive development, learning, and developmental research. *New directions for child and adolescent development*, 2013(139), 1–9. <https://doi.org/10.1002/cad.20026>
- Böhme, R., Munser-Kiefer, M. & Prestridge, S. (2020). Lernunterstützung mit digitalen Medien in der Grundschule. *Zeitschrift für Grundschulforschung*, 13(1), 1–14. <https://doi.org/10.1007/s42278-019-00066-3>
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons. <https://doi.org/10.1002/9781118619179>
- Bollen, K. A. & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, 21(2), 205–229.
- Bong, M. & Skaalvik, E. M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational psychology review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>
- Bortz, J. & Döring, N. (2016). *Forschungsmethoden und Evaluation: In den Sozial- und Humanwissenschaften* (5. Aufl.). Springer-Verlag. http://deposit.d-nb.de/cgi-bin/dokserv?id=2841557&prov=M&dok_var=1&dok_ext=htm
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Springer. <https://doi.org/10.1007/978-3-642-12770-0>
- Boruch, R. F. (1997). *Randomized experiments for planning and evaluation: A practical guide*. Sage.
- Bos, W., Eickelmann, B., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., Schulz-Zander, R. & Wendt, H. (2014). *ICILS 2013: Computer-und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich*. Waxmann.
- Boyle, E. A., Connolly, T. M. & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2(2), 69–74. <https://doi.org/10.1016/j.entcom.2010.12.002>
- Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C. & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, 94, 178–192. <https://doi.org/10.1016/j.compedu.2015.11.003>
- Bracken, B. A. (1996). *Handbook of self-concept: Developmental, social, and clinical considerations*. Wiley series on personality processes. J. Wiley.
- Braghirolli, L. F., Ribeiro, J. L. D., Weise, A. D. & Pizzolato, M. (2016). Benefits of educational games as an introductory activity in industrial engineering education. *Computers in Human Behavior*, 58, 315–324. <https://doi.org/10.1016/j.chb.2015.12.063>
- Brähler, E., Holling, H., Leutner, D. & Petermann, F. (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Tests*. Hogrefe.
- Brauner, P. & Ziefle, M. (2022). Beyond playful learning – Serious games for the human-centric digital transformation of production and a design process model. *Technology in Society*, 71, 102140. <https://doi.org/10.1016/j.techsoc.2022.102140>
- Brennan, R. L. (2001). An essay on the history and future of reliability from the perspective of replications. *Journal of Educational Measurement*, 38(4), 295–317.
- Bressler, D. M. & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. *Journal of Computer Assisted Learning*, 29(6), 505–517. <https://doi.org/10.1111/jcal.12008>
- Bressler, D. M. & Bodzin, A. M. (2016). Investigating Flow Experience and Scientific Practices During a Mobile Serious Educational Game. *Journal of Science Education and Technology*, 25(5), 795–805. <https://doi.org/10.1007/s10956-016-9639-z>

- Bressler, D. M., Bodzin, A. M. & Tutwiler, M. S. (2019). Engaging middle school students in scientific practice with a collaborative mobile game. *Journal of Computer Assisted Learning*, 35(2), 197–207. <https://doi.org/10.1111/jcal.12321>
- Breuer, J. & Bente, G. (2010). Why so serious? On the relation of serious games and learning. *Journal for Computer Game Culture*, 4(1), 7–24.
- Breusch, T. S. & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5), 1287. <https://doi.org/10.2307/1911963>
- Brewer, B. W., van Raalte, J. L., Linder, D. E. & van Raalte, N. S. (1991). Peak performance and the perils of retrospective introspection. *Journal of sport and exercise psychology*, 13(3), 227–238.
- Brom, C., Preuss, M. & Klement, D. (2011). Are educational computer micro-games engaging and effective for knowledge acquisition at high-schools? A quasi-experimental study. *Computers & Education*, 57(3), 1971–1988. <https://doi.org/10.1016/j.compedu.2011.04.007>
- Broos, A. & Roe, K. (2006). The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents. *Poetics*, 34(4-5), 306–317. <https://doi.org/10.1016/j.poetic.2006.05.002>
- Brown, G. T. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological reports*, 99(1), 166–170.
- Brown, R. M., Hall, L. R., Holtzer, R., Brown, S. L. & Brown, N. L. (1997). Gender and video game performance. *Sex roles*, 36, 793–812.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 136–162). SAGE Publications.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31, 21–32.
- Bühner, M. (2011). *Einführung in die Test-und Fragebogenkonstruktion*. Pearson Deutschland GmbH.
- Burde, J.-P. (2018). *Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells*. Logos Verlag. <https://doi.org/10.30819/4726>
- Bürger, N., Haselmann, S., Baumgart, J., Prinz, G., Girnat, B., Meisert, A., Menthe, J., Schmidt-Thieme, B. & Wecker, C. (2021). Jenseits von Professionswissen: Eine systematische Überblicksarbeit zu einstellungs- und motivationsbezogenen Einflussfaktoren auf die Nutzung digitaler Technologien im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 24(5), 1087–1112. <https://doi.org/10.1007/s11618-021-01050-3>
- Burgers, C., Eden, A., van Engelenburg, M. D. & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior*, 48, 94–103. <https://doi.org/10.1016/j.chb.2015.01.038>
- Burzin, S., Kahnt, M., Moussa, A., Müller, W., Nawrath, D. & Rutscher, C. (2016). *Fokus Physik - Baden-Württemberg* (1. Auflage, 1. Druck). *Fokus Physik - Neubearbeitung - Gymnasium Baden-Württemberg*. Cornelsen.
- Byrne, B. M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. 04715993.
- Caillois, R. (2001). *Man, play, and games*. University of Illinois press.
- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of personality and social psychology*, 59(3), 538–549. <https://doi.org/10.1037/0022-3514.59.3.538>
- Capello, D. (2023, 15. Juli). *Aseprite*. <https://www.aseprite.org/>
- Carmesin, H.-O., Kahle, J., Kienle, R., Konrad, U., Küblbeck, J., Pardall, C.-J., Rager, B. & Trumme, T. (2016). *Universum Physik - Berlin/Brandenburg* (1. Auflage, 1. Druck). Cornelsen.
- Cassell, J. & Jenkins, H. (1998). Chess for girls? Feminism and computer games. In *From Barbie to Mortal Kombat: gender and computer games* (S. 2–45).

- Chambers, J., Eddy, W., Härdle, W., Sheather, S., Tierney, L., Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer New York. <https://doi.org/10.1007/978-0-387-21706-2>
- Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of clinical and experimental neuropsychology*, 21(4), 559–566. <https://doi.org/10.1076/jcen.21.4.559.889>
- Chen, C.-H. (2019). The impacts of peer competition-based science gameplay on conceptual knowledge, intrinsic motivation, and learning behavioral patterns. *Educational Technology Research and Development*, 67(1), 179–198. <https://doi.org/10.1007/s11423-018-9635-5>
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), 31–34. <https://doi.org/10.1145/1232743.1232769>
- Chen, L. H., Ye, Y.-C., Chen, M.-Y. & Tung, I.-W. (2010). Alegría! Flow in leisure and life satisfaction: The mediating role of event satisfaction using data from an acrobatics show. *Social indicators research*, 99, 301–313.
- Cheng, M.-T. & Annetta, L. (2012). Students' learning outcomes and learning experiences through playing a Serious Educational Game. *Journal of Biological Education*, 46(4), 203–213. <https://doi.org/10.1080/00219266.2012.688848>
- Cheng, M.-T., Chen, J.-H., Chu, S.-J. & Chen, S.-Y. (2015). The use of serious games in science education: a review of selected empirical research from 2002 to 2013. *Journal of Computers in Education*, 2(3), 353–375. <https://doi.org/10.1007/s40692-015-0039-9>
- Cheng, M.-T., She, H.-C. & Annetta, L. A. (2015). Game immersion experience: its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning*, 31(3), 232–253. <https://doi.org/10.1111/jcal.12066>
- Cheng, M.-T., Su, T., Huang, W.-Y. & Chen, J.-H. (2014). An educational game for learning human immunology: What do students learn and how do they perceive? *British Journal of Educational Technology*, 45(5), 820–833. <https://doi.org/10.1111/bjet.12098>
- Clark, D. B., Tanner-Smith, E. E. & Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of educational research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Closset, J. L. (1984). Woher stammen bestimmte „Fehler“ von Schülern und Studenten aus dem Bereich der Elektrizitätslehre? Kann man sie beheben. *Der Physikunterricht*, 18(2), 21–31.
- Cohen, J. (1983). The cost of dichotomization. *Applied psychological measurement*, 7(3), 249–253.
- Cohen, J. (1988). Set Correlation and Contingency Tables. *Applied psychological measurement*, 12(4), 425–434. <https://doi.org/10.1177/014662168801200410>
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Coleman, J. M. & Fults, B. A. (1982). Self-Concept and the Gifted Classroom: The Role of Social Comparisons. *Gifted Child Quarterly*, 26(3), 116–120. <https://doi.org/10.1177/001698628202600305>
- Colquitt, J. A., LePine, J. A. & Noe, R. A. (2000). Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research. *Journal of applied psychology*, 85(5), 678.

- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T. & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661–686. <https://doi.org/10.1016/j.compedu.2012.03.004>
- Cook, R. D. & Weisberg, S. (1995). *Residuals and influence in regression* (Reprint). *Monographs on statistics and applied probability: Bd. 18*. Chapman & Hall.
- Couceiro, R. M., Papastergiou, M., Kordaki, M. & Veloso, A. I. (2013). Design and evaluation of a computer game for the learning of Information and Communication Technologies (ICT) concepts by physical education and sport science students. *Education and Information Technologies*, 18, 531–554.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety* (1. Aufl.). *The Jossey-Bass behavioral science series*. Jossey-Bass.
- Csikszentmihalyi, M. (2014). *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*. Springer Netherlands. <https://doi.org/10.1007/978-94-017-9088-8>
- Csikszentmihalyi, M. & Csikszentmihályi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge University Press.
- Csikszentmihalyi, M. & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of personality and social psychology*, 56(5), 815–822. <https://doi.org/10.1037/0022-3514.56.5.815>
- Csikszentmihalyi, M. & Schiefele, U. (1993). Die Qualität des Erlebens und der Prozess des Lernens. *Zeitschrift für Pädagogik*, 39(2), 207–221. <https://doi.org/10.25656/01:11172>
- D'Agostino, R. B. & Belanger, A. (1990). A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*, 44(4), 316. <https://doi.org/10.2307/2684359>
- Davidson, R. & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US. <https://doi.org/10.1007/978-1-4899-2271-7>
- Deci, E. L. & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Dickey, M. D. (2006). Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational Technology Research and Development*, 54(3), 245–263. <https://doi.org/10.1007/S11423-006-8806-Y>
- Dickey, M. D. (2011). Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. *British Journal of Educational Technology*, 42(3), 456–469. <https://doi.org/10.1111/j.1467-8535.2009.01032.x>
- Dilling, F. (2022). Konstruktivistische Lerntheorie. *Begründungsprozesse im Kontext von (digitalen) Medien im Mathematikunterricht: Wissensentwicklung auf der Grundlage empirischer Settings*, 9–35.
- Djaouti, D., Alvarez, J. & Jessel, J.-P. (2011). Classifying Serious Games: The G/P/S model. In P. Felicia (Hrsg.), *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches* (S. 118–136). Information Science Reference. <https://doi.org/10.4018/978-1-60960-495-0.ch006>
- Dopatka, L., Spatz, V., Burde, J.-P., Wilhelm, T., Ivanjek, L., Hopf, M., Haagen-Schützenhöfer, C. & Schubatzky, T. (2018). Kontexte in der Elektrizitätslehre im Rahmen des Projekts EPO-EKo. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe: Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Kiel 2018* (217-220).

- Draper, N. R. & Smith, H. (1998). *Applied regression analysis* (3rd ed.). *Wiley series in probability and statistics*. Wiley. <http://www.loc.gov/catdir/bios/wiley042/97017969.html>
- Durbin, J. & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, 37(3/4), 409–428.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040.
- Eddy, S. L. & Brownell, S. E. (2016). Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, 12(2), Artikel 020106. <https://doi.org/10.1103/PhysRevPhysEducRes.12.020106>
- Egbert, J. (2004). A study of flow theory in the foreign language classroom. *Canadian modern language review*, 60(5), 549–586.
- Eickelmann, B., Bos, W. & Labusch, A. (2019). *Die Studie ICILs 2018 im Überblick. Zentrale Ergebnisse und mögliche Entwicklungsperspektiven*. Waxmann. https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM_2019.pdf <https://doi.org/10.25656/01:18319>
- Eid, M., Gollwitzer, M. & Schmitt, M. (2013). *Statistik und Forschungsmethoden: Lehrbuch. Mit Online-Materialien* (Deutsche Erstausgabe, 3., korrigierte Aufl.). Beltz.
- Eid, M. & Schmidt, K. (2014). *Testtheorie und Testkonstruktion*. Hogrefe Verlag GmbH & Company KG.
- Engeser, S. & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and emotion*, 32(3), 158–172.
- Engeser, S., Rheinberg, F., Vollmeyer, R. & Bischoff, J. (2005). Motivation, Flow-Erleben und Lernerleistung in universitären Lernsettings. *Zeitschrift für Pädagogische Psychologie*, 19(3), 159–172. <https://doi.org/10.1024/1010-0652.19.3.159>
- Engeser, S., Schiepe-Tiska, A. & Peifer, C. (2021). Historical lines and an overview of current research on flow. *Advances in flow research*, 1–29.
- Engeser, S. & Vollmeyer, R. (2005). Tätigkeitsanreize und Flow-Erleben. In R. Vollmeyer & J. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendung* (S. 59–71). Kohlhammer.
- Ertl, B. (2016). Game-Based Learning for Supporting Self-Confidence and Motivation of Female STEM Students. In L.-J. Thoms & R. Girwidz (Hrsg.), *Selected Papers from the 20th International Conference on Multimedia in Physics Teaching and Learning* (S. 129–137). Ludwig-Maximilians-Universität München.
- Evans, R. I. (1989). *Albert Bandura: The man and his ideas—a dialogue*. Praeger Publishers.
- Eyupoglu, T. F. & Nietfeld, J. L. (2019). Intrinsic motivation in game-based learning environments. In D. Ifenthaler & Y. J. Kim (Hrsg.), *Game-based assessment revisited: Advances in Game-Based Learning* (S. 85–102). Springer International Publishing. https://doi.org/10.1007/978-3-030-15569-8_5
- Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, 41(4), 1149–1160.
- Feierabend, S., Rathgeb, T., Kheredmand, H. & Glöckler, S. (2020). *JIM-Studie 2020: Jugend, Information, Medien: Basisuntersuchung zum Medienumgang 12-bis 19-Jähriger*. Medienpädagogischer Forschungsverbund Südwest (mpfs).
- Ferla, J., Valcke, M. & Cai, Y [Yonghong] (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and individual differences*, 19(4), 499–505. <https://doi.org/10.1016/j.lindif.2009.05.004>
- Field, Z., Miles, J. & Field, A. (2012). Discovering statistics using R. *Discovering Statistics Using R*, 1–992.

- Fischer, N., Pfeiffer, T. & Dickhäuser, O. (2021). Die Bedeutung von Motivation für Leistung. In N. Fischer, T. Pfeiffer & O. Dickhäuser (Hrsg.), *Stark im Scheitern - Motivation nach Misserfolgen* (S. 3–10). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-33281-5_1
- Fournier, J., Gaudreau, P., Demontrond-Behr, P., Visioli, J., Forest, J. & Jackson, S. (2007). French translation of the Flow State Scale-2: Factor structure, cross-cultural invariance, and associations with goal attainment. *Psychology of Sport and Exercise*, 8(6), 897–916. <https://doi.org/10.1016/j.psychsport.2006.07.007>
- Franz, V. S. T. (2020). *Zum Zusammenhang von Leistungsmotivation, Flow-Erleben und subjektivem Wohlbefinden: Eine empirische Studie an Schülerinnen und Schülern der Primar- und Sekundarstufe* [Dissertation Philosophie]. Universität Trier. https://ubt.opus.hbz-nrw.de/files/1465/Dissertation_Viktoria+S.+Franz.pdf <https://doi.org/10.25353/ubtr-xxxx-c72a-d7c0>
- Fritz, C. O., Morris, P. E. & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology. General*, 141(1), 2–18. <https://doi.org/10.1037/a0024338>
- Furió, D., González-Gancedo, S., Juan, M.-C., Seguí, I. & Rando, N. (2013). Evaluation of learning outcomes using an educational iPhone game vs. traditional game. *Computers & Education*, 64, 1–23.
- Ganguin, S. (2010). *Computerspiele und lebenslanges Lernen: Eine Synthese von Gegensätzen* (1. Aufl.). *Medienbildung und Gesellschaft: Bd. 13*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-92433-5>
- Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, 33(4), 441–467. <https://doi.org/10.1177/1046878102238607>
- Gastwirth, J. L., Gel, Y. R., Hui, W. W., Lyubchich, V., Miao, W. & Noguchi, K. (2017). lawstat: Tools for biostatistics, public policy, and law. *R package version*, 3.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy* (1st ed.). Palgrave Macmillan. <https://doi.org/10.1145/950566.950595>
- Gee, J. P. (2007). *Good video games + good learning: Collected essays on video games, learning, and literacy. New literacies and digital epistemologies: v. 27*. Peter Lang US. <https://doi.org/10.3726/978-1-4539-1162-4>
- Giammarco, E. A., Schneider, T. J., Carswell, J. J. & Knipe, W. S. (2015). Video game preferences and their relation to career interests. *Personality and Individual Differences*, 73, 98–104. <https://doi.org/10.1016/j.paid.2014.09.036>
- Giessen, H. W. (2015). Serious Games Effects: An Overview. *Procedia - Social and Behavioral Sciences*, 174, 2240–2244. <https://doi.org/10.1016/j.sbspro.2015.01.881>
- GIMP. (2023, 11. Juli). *GIMP*. <https://www.gimp.org/>
- Gliner, J. A., Morgan, G. A. & Harmon, R. J. (2003). Pretest-posttest comparison group designs: analysis and interpretation. *Journal of the American Academy of Child & Adolescent Psychiatry*.
- Gnauk, B., Dannecker, L. & Hahmann, M. (2012). Leveraging gamification in demand dispatch systems. In D. Srivastava & I. Ari (Hrsg.), *Proceedings of the 2012 Joint EDBT/ICDT Workshops* (S. 103–110). ACM. <https://doi.org/10.1145/2320765.2320799>
- Godot. (2023, 12. Juli). *Godot Engine - Free and open source 2D and 3D game engine*. <https://godotengine.org/>
- Gredler, M. E. (1996). 17. Educational games and simulations: A technology in search of a (research) paradigm. *Technology*, 39, 521–540.

- Gronier, G. & Baudet, A. (2019). Does Progress Bars' Behavior Influence the User Experience in Human-Computer Interaction? *Psychology and Cognitive Sciences Open Journal*, 4(2), 57–64.
- Groos, K. (1899). *Die Spiele der Menschen*. Fischer.
- Hamari, J. & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale-2. *Computers in Human Behavior*, 40, 133–143. <https://doi.org/10.1016/j.chb.2014.07.048>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J. & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Hardin, A. M., Looney, C. A. & Fuller, M. A. (2014). Self-efficacy, learning method appropriation and software skills acquisition in learner-controlled CSSTS environments. *Information Systems Journal*, 24(1), 3–27.
- Harter, S. (1988). Causes, correlates and the functional role of global self-worth. *Perceptions of Competence and incompetence across the life-span*.
- Härtig, H., Ostermann, A., Ropohl, M., Schwanewedel, J., Kampschulte, L. & Lindmeier, A. (2021). Gibt es einen fachspezifischen Medieneinsatz im naturwissenschaftlichen Fachunterricht? Ergebnisse einer Fragebogenerhebung. *Zeitschrift für Didaktik der Naturwissenschaften*, 27(1), 139–154. <https://doi.org/10.1007/s40573-021-00130-5>
- Hastie, T., Friedman, J. & Tibshirani, R. (2001). *The Elements of Statistical Learning*. Springer. <https://doi.org/10.1007/978-0-387-21606-5>
- Hattie, J. (1992). *Self-concept*. L. Earlbaum Associates.
- Hattie, J. (1999). Influences on student learning. *Inaugural lecture given on August, 2(1999)*, 21.
- Häußler, P. & Hoffmann, L. (1995). Physikunterricht-an den Interessen von Mädchen und Jungen orientiert. *Unterrichtswissenschaft*, 23(2), 107–126.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- Heckhausen, H. (1969). Förderung der Lernmotivierung und der intellektuellen Tüchtigkeiten. *Begabung und Lernen*, 2, 193–228.
- Heckhausen, J. & Heckhausen, H. (Hrsg.). (2006). *Springer-Lehrbuch. Motivation und Handeln*. Springer-Verlag. <https://doi.org/10.1007/3-540-29975-0>
- Hellmich, F. & Günther, F. (2011). Entwicklung von Selbstkonzepten bei Kindern im Grundschulalter–ein Überblick. In F. Hellmich (Hrsg.), *Selbstkonzepte im Grundschulalter: Modelle, empirische Ergebnisse, pädagogische Konsequenzen* (S. 19–46). Kohlhammer Verlag.
- Hoblitz, A. (2015). *Spielend Lernen im Flow: Die motivationale Wirkung von Serious Games im Schulunterricht*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-11376-6>
- Hoffmann, L., Häußler, P. & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik*. IPN.
- Holland, W., Jenkins, H., Squire, K. D., Miller, Z. N., O'Driscoll, A., Shresthova, S., Soley, J., Targum, E., Yew, P. T. B. & Todd, K. (2003). Theory by design. *The video game theory reader*, 25–46.
- Hoyle, R. H. (1999). *Statistical strategies for small sample research*. Sage.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huizinga, J. (1938). *Homo Ludens: Vom Ursprung der Kultur im Spiel* (H. Nachod, Übers.) (23. Auflage 2013). *rororo Rowohlts Enzyklopädie: Bd. 55435*. Rowohlt Taschenbuch Verlag.
- Hung, C.-Y., Sun, J. C.-Y. & Yu, P.-T. (2015). The benefits of a challenge: student motivation and flow experience in tablet-PC-game-based learning. *Interactive Learning Environments*, 23(2), 172–190. <https://doi.org/10.1080/10494820.2014.997248>

- Inal, Y. & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology*, 38(3), 455–464. <https://doi.org/10.1111/j.1467-8535.2007.00709.x>
- Janneck, M., Vincent-Höper, S. & Ehrhardt, J. (2012). Das computerbezogene Selbstkonzept: Eine gender-sensitive Studie. In H. Reiterer & O. Deussen (Hrsg.), *Mensch & Computer 2012: 12. fachübergreifende Konferenz für interaktive und kooperative Medien*. (S. 243–252). Oldenbourg Wissenschaftsverlag. <https://doi.org/10.1524/9783486718782.243>
- Jarvis, S. & Freitas, S. de (2009). Evaluation of an Immersive Learning Programme to Support Triage Training. In *2009 Conference in Games and Virtual Worlds for Serious Applications* (S. 117–122). IEEE. <https://doi.org/10.1109/VS-GAMES.2009.31>
- Jerusalem, M. & Satow, L. (2009). Schulbezogene Selbstwirksamkeit. In *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* : Humboldt Universität.
- Jones, M. G. (1998). Creating Electronic Learning Environments: Games, Flow, and the User Interface. In N. J. Maushak, C. Schlosser, T. N. Lloyd & M. Simonson (Hrsg.), *Proceedings of Selected Research and Development Presentations at the National Convention of the Association for Educational Communications and Technology* (S. 205–214).
- Jonkisz, E., Moosbrugger, H. & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. *Testtheorie und Fragebogenkonstruktion*, 27–74.
- Kapp, F., Spangenberg, P., Kruse, L. & Narciss, S. (2019). Investigating changes in self-evaluation of technical competences in the serious game Serena Supergreen: Findings, challenges and lessons learned. *Metacognition and Learning*, 14(3), 387–411. <https://doi.org/10.1007/s11409-019-09209-4>
- Ke, F. (2009). A qualitative meta-analysis of computer games as learning tools. *Handbook of research on effective electronic gaming in education*, 1–32.
- Ke, F. & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology*, 44(2), 225–242. <https://doi.org/10.1111/j.1467-8535.2012.01326.x>
- Kebritchi, M. & Hirumi, A. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers & Education*, 51(4), 1729–1743. <https://doi.org/10.1016/j.compedu.2008.05.004>
- Kerres, M., Bormann, M. & Vervenne, M. (2009). Didaktische Konzeption von Serious Games: Zur Verknüpfung von Spiel- und Lernangeboten. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 1–16. <https://doi.org/10.21240/mpaed/00/2009.08.25.X>
- Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 99–111. <https://doi.org/10.1007/s10956-006-9038-y>
- Kickmeier-Rust, M. D., Schwarz, D., Albert, D., Verpoorten, D., Jean-Loup Castaigne & Bopp, M. (2006). *The Elektra Project: Towards a New Learning Experience: Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society*. <https://doi.org/10.13140/2.1.2272.8646>
- Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and Higher Education*, 8(1), 13–24. <https://doi.org/10.1016/j.iheduc.2004.12.001>
- Kiili, K., Freitas, S. de, Arnab, S. & Lainema, T. (2012). The Design Principles for Flow Experience in Educational Games. *Procedia Computer Science*, 15, 78–91. <https://doi.org/10.1016/j.procs.2012.10.060>

- Kilian, U. (Hrsg.). (2017). *Dorn Bader - Physik Gymnasium Baden-Württemberg*. Schroedel Westermann.
- Klassen, R. M. & Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, 12, 59–76. <https://doi.org/10.1016/j.edurev.2014.06.001>
- Kleickmann, T., Brehl, T., Saß, S., Prenzel, M. & Köller, O. (2012). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 123–169). Waxmann.
- Klimmt, C. (2009). Serious games and social change: Why they (should) work. In U. Ritterfeld, M. J. Cody & P. Vorderer (Hrsg.), *Serious Games: Mechanisms and Effects* (S. 270–292). Routledge.
- Klimmt, C. & Hartmann, T. (2006). Effectance, Self-Efficacy, and the Motivation to Play Video Games. In P. Vorderer & J. Bryant (Hrsg.), *Playing Video Games: Motives, responses, and consequences: Motives, Responses, and Consequences* (S. 153–168). Lawrence Erlbaum Associates Publishers.
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. SAGE Publications.
- KMK. (2004). *Beschlüsse der Kultusministerkonferenz*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Physik-Mittleren-SA.pdf
- KMK. (2016). *Strategie der Kultusministerkonferenz: "Bildung in der digitalen Welt": Beschluss der Kultusministerkonferenz vom 08.12.2016 in der Fassung vom 07.12.2017*.
- Kolen, M. J. & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking: Methods and Practices*. Springer. <https://doi.org/10.1007/978-1-4939-0317-7>
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 27–39. <https://doi.org/10.1024/1010-0652.20.12.27>
- Komsta, L. & Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and related tests. *R package version*, 14(1).
- König, J., Doll, J., Buchholtz, N., Förster, S., Kaspar, K., Rühl, A. M., ... & Kaiser, G. (2017). Pädagogisches Wissen versus fachdidaktisches Wissen? Struktur des professionellen Wissens bei angehenden Deutsch-, Englisch- und Mathematiklehrkräften im Studium. *Zeitschrift für Erziehungswissenschaft*.
- Kopp, J. & Lois, D. (2014). *Sozialwissenschaftliche Datenanalyse: Eine Einführung*. Springer-Verlag.
- Kosuch, R. (2010). Selbstwirksamkeit und Geschlecht–Impulse für die MINT-Didaktik. *«Gender und MINT» Schlussfolgerungen für Unterricht, Beruf und Studium*, 12.
- Kraiger, K., Ford, J. K. & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, 78(2), 311.
- Krath, J., Schürmann, L. & Korflesch, H. F. von (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, 125, 106963. <https://doi.org/10.1016/j.chb.2021.106963>
- Kromrey, J. D. & Foster-Johnson, L. (1998). Mean Centering in Moderated Multiple Regression: Much Ado about Nothing. *Educational and Psychological Measurement*, 58(1), 42–67. <https://doi.org/10.1177/0013164498058001005>

- Kubinger, K. D., Rasch, D. & Moder, K. (2009). Zur Legende der Voraussetzungen des t-Tests für unabhängige Stichproben. *Psychologische Rundschau*, 60(1), 26–27.
- Kutner, M. H., Nachtsheim, C. & Neter, J. (2004). *Applied linear regression models* (4. ed., internat. ed.). *The MacGraw-Hill/Irwin series*. McGraw-Hill/Irwin.
<http://www.loc.gov/catdir/description/mh031/2003044224.html>
- Lamb, R. L., Annetta, L. A., Firestone, J. & Etopio, E. (2018). A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Computers in Human Behavior*, 80, 158–167. <https://doi.org/10.1016/j.chb.2017.10.040>
- Landhäuser, A. & Keller, J. (2012). Flow and its affective, cognitive, and performance-related consequences. In S. Engeser (Hrsg.), *Advances in Flow Research* (S. 65–85). Springer US.
- Lawless, K. A. & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, 25, 117–131.
- Lazarus, M. (1883). *Über die Reize des Spiels*. Dümmler.
- Le, S., Weber, P. & Ebner, M. (2013). Game-Based Learning - Spielend Lernen? In S. Schön & M. Ebner (Hrsg.), *Lehrbuch für Lernen und Lehren mit Technologien* (2. Aufl., S. 267–276). epubli.
- Lechte, M.-A. (2008). *Sinnbezüge, Interesse und Physik: eine empirische Untersuchung zum Erleben von Physik aus Sicht von Schülerinnen und Schülern*. Verlag Barbara Budrich.
- Lee, K. M. (2000). MUD and Self Efficacy. *Educational Media International*, 37(3), 177–183. <https://doi.org/10.1080/09523980050184745>
- Leiner, D. J. (2023). *SoSci Survey* (Version 3.4.19) [Computer software]. <https://www.soscisurvey.de>
- Leisen, J. (2005). Muss ich jetzt auch noch Sprache unterrichten? Sprache und Physikunterricht. *Unterricht Physik*, 16(3), 4–9.
- Lengning, A. (2009). Neugier und Exploration. In V. Brandstätter & J. H. Otto (Hrsg.), *Handbuch der Psychologie: Bd. 11. Handbuch der Allgemeinen Psychologie - Motivation und Emotion* (S. 252–257). Hogrefe.
- Leung, L. (2020). Exploring the relationship between smartphone activities, flow experience, and boredom in free time. *Computers in Human Behavior*, 103, 130–139.
- Liao, C.-W., Chen, C.-H. & Shih, S.-J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. *Computers & Education*, 133(2), 43–55.
- Lieberman, D. A. (2006). What can we learn from playing interactive games. In P. Vorderer & J. Bryant (Hrsg.), *Playing Video Games: Motives, responses, and consequences: Motives, Responses, and Consequences* (S. 379–397). Lawrence Erlbaum Associates Publishers.
- Locke, E. A. & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.
- Løvoll, H. S. & Vittersø, J. (2014). Can balance be boring? A critique of the “challenges should match skills” hypotheses in flow theory. *Social indicators research*, 115, 117–136.
- Lu, Y.-L. & Lien, C.-J. (2020). Are they learning or playing? Students’ perception traits and their learning self-efficacy in a game-based learning environment. *Journal of Educational Computing Research*, 57(8), 1879–1909. <https://doi.org/10.1177/0735633118820684>
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive science*, 5(4), 333–369.
- Marr, A. C. (2010). *Serious Games für die Informations-und Wissensvermittlung-Bibliotheken auf neuen Wegen*. BIT Verlag.
- Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, 82(4), 623.

- Marsh, H. W., Byrne, B. M. & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80*(3), 366.
- Marsh, H. W. & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on psychological science : a journal of the Association for Psychological Science, 1*(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>
- Marsh, H. W. & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. *04715993*.
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T. & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331.
- Massimini, F. & Carli, M. (1988). 16. The systematic assessment of flow in daily experience.
- Mayer, I., Bekebrede, G., Harteveld, C., Warmelink, H., Zhou, Q., van Ruijven, T., Lo, J., Kortmann, R. & Wenzler, I. (2014). The research and evaluation of serious games: Toward a comprehensive methodology. *British Journal of Educational Technology, 45*(3), 502–527. <https://doi.org/10.1111/bjet.12067>
- Mayer, R. (2019). Computer Games in Education. *Annual Review of Psychology, 70*(1), 531–549. <https://doi.org/10.1146/annurev-psych-010418-102744>
- Mäyrä, F. & Ermi, L. (2011). Fundamental components of the gameplay experience. In S. Günzel, M. Liebe & D. Mersch (Hrsg.), *DIGAREC Keynote-Lectures 2009/10* (S. 88–115). University Press.
- McCowan, R. J. & McCowan, S. C. (1999). Item Analysis for Criterion-Referenced Tests. *Online Submission*.
- McDermott, L. C. & Shaffer, P. S. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics, 60*(11), 994–1003. <https://doi.org/10.1119/1.17003>
- Michael, D. & Chen, S. (Hrsg.). (2006). *Serious Games : Games That Educate, Train, and Inform*. Thomson Course Technology PTR.
- Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and immersion in video games: The aftermath of a conceptual challenge. *Frontiers in psychology, 9*, 393107.
- Mitra, B. M. & Golz, P. (2016). Exploring intrinsic gender identity using second life. *Journal For Virtual Worlds Research, 9*(2).
- MKJS (Hrsg.). (2015a). *Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2015): Bildungsplan 2016. Allgemeinbildende Schulen. Gymnasium. Physik*. https://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_PH.pdf
- MKJS (Hrsg.). (2015b). *Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2015): Bildungsplan 2016. Gemeinsamer Bildungsplan der Sekundarstufe 1. Physik*. http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_PH.pdf
- MKJS. (2020). *Verwaltungsvorschrift des Kultusministeriums zur Umsetzung der Verwaltungsvereinbarung DigitalPakt Schule 2019 bis 2024*. Ministerium für Kultus, Jugend und Sport Baden-Württemberg. https://km-bw.de/site/pbs-bw-km-root/get/documents_E-1891332534/KULTUS.Dachmandant/KULTUS/KM-Homepage/Artikelseiten%20KP-KM/Digitalisierung/Digitalpakt%20Schule/22-12-13%20Verwaltungsvereinbarung%20DigitalPakt%20Schule%202019-2024_bf.pdf
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte. *Psychologische Rundschau, 55*(1), 19–27.

- Möller, J. & Trautwein, U. (2009). Selbstkonzept. *Pädagogische psychologie*, 179–203.
- Mondak, J. J. & Davis, B. C. (2001). Asked and answered: Knowledge levels when we will not take “don't know” for an answer. *Political Behavior*, 23, 199–224.
- Moneta, G. B. (2021). On the conceptualization and measurement of flow. In C. Peifer & S. Engeser (Hrsg.), *Advances in Flow Research* (S. 31–69). Springer; Springer International Publishing.
- Moneta, G. B. & Csikszentmihalyi, M. (1999). Models of concentration in natural environments: A comparative approach based on streams of experiential data. *Social Behavior and Personality: an international journal*, 27(6), 603–637.
<https://doi.org/10.2224/sbp.1999.27.6.603>
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2020). *Testtheorie und Fragebogenkonstruktion*. Springer.
- Moreno, R. & Mayer, R. (2007). Interactive multimodal learning environments. *Educational psychology review*, 19(3), 309–326.
- Mummendey, H. D. (2006). *Psychologie des "Selbst": Theorien, Methoden und Ergebnisse der Selbstkonzeptforschung*. Hogrefe.
- Muthén, L. K. & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620.
- Nacke, L. E. & Lindley, C. A. (2009). Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience. *Loading...The Journal of the Canadian Game Studies Association*, 3(5), 1–21.
- Nelson, B. C. & Ketelhut, D. J. (2008). Exploring embedded guidance and self-efficacy in educational multi-user virtual environments. *International Journal of Computer-Supported Collaborative Learning*, 3, 413–427.
- Nelson, L. J. & Cooper, J. (1997). Gender differences in children's reactions to success and failure with computers. *Computers in Human Behavior*, 13(2), 247–267.
[https://doi.org/10.1016/S0747-5632\(97\)00008-3](https://doi.org/10.1016/S0747-5632(97)00008-3)
- Nissen, J. M. (2019). Gender differences in self-efficacy states in high school physics. *Physical Review Physics Education Research*, 15(1), Artikel 013102.
- Nordin, A. I., Ali, J., Animashaun, A., Asch, J., Adams, J., & Cairns, P. (2013). Attention, time perception and immersion in games. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1089-1094).
- Nunnally, J. C. & Bernstein, I. H. (2008). *Psychometric theory* (3. ed. [Nachdr.]. *McGraw-Hill series in psychology*. McGraw-Hill.
- O'Reilly, T. & McNamara, D. S. (2007). The Impact of Science Knowledge, Reading Skill, and Reading Strategy Knowledge on More Traditional “High-Stakes” Measures of High School Students’ Science Achievement. *American Educational Research Journal*, 44(1), 161–196.
<https://doi.org/10.3102/0002831206298171>
- O'Brien, R. G. (1978). Robust techniques for testing heterogeneity of variance effects in factorial designs. *Psychometrika*, 43(3), 327–342. <https://doi.org/10.1007/BF02293643>
- OECD. (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school. Programme for International Student Assessment*. OECD.
- OECD. (2016). *PISA 2015 Ergebnisse: Exzellenz und Chancengerechtigkeit in der Bildung*. Bertelsmann. <https://doi.org/10.3278/6004573w>
- Oerter, R. (2011). *Psychologie des Spiels: Ein handlungstheoretischer Ansatz* (Durchges. Neuausg., 2. Aufl.). *Beltz-Taschenbuch Psychologie: Bd. 46*. Beltz.
- O'Neil, H. F., Wainess, R. & Baker, E. L. (2005). Classification of learning outcomes: evidence from the computer games literature. *The Curriculum Journal*, 16(4), 455–474.
<https://doi.org/10.1080/09585170500384529>

- Orvis, K. A., Horn, D. B. & Belanich, J. (2008). The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Computers in Human Behavior*, 24(5), 2415–2433.
- Pajares, F. (1997). Current directions in self-efficacy research. *Advances in motivation and achievement*, 10(149), 1–49.
- Pajares, F. & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86(2), 193–203. <https://doi.org/10.1037/0022-0663.86.2.193>
- Pajares, F. & Schunk, D. H. (2006). Self-efficacy and self-concept beliefs: Jointly Contributing To The Quality Of Human Life. In H. W. Marsh, R. G. Craven & D. M. McInerney (Hrsg.), *New Frontiers for Self Research* (S. 95–122). Information Age Publishing.
- Paraskeva, F., Mysirlaki, S. & Papagianni, A. (2010). Multiplayer online games as educational tools: Facing new challenges in learning. *Computers & Education*, 54(2), 498–505.
- Pavlas, D. (2010). *A Model Of Flow And Play In Game-based Learning: The Impact of Game Characteristics, Player Traits and Player States* [Dissertation, University of Central Florida, Orlando, FL]. RIS.
- Pavlas, D., Heyne, K., Bedwell, W. L., Lazzara, E. H. & Salas, E. (2010). Game-based Learning: The Impact of Flow State and Videogame Self-efficacy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 54(28), 2398–2402. <https://doi.org/10.1177/154193121005402808>
- Peifer, C. (2017). Zum Zusammenhang zwischen Flow-Erleben und Stress im Kontext von Leistung und Wohlbefinden. *Zum Stand Positiv-Psychologischer Forschung im deutschsprachigen Raum–State of the Art*, 18–36.
- Peifer, C., Schönfeld, P., Wolters, G., Aust, F. & Margraf, J. (2020). Well Done! Effects of Positive Feedback on Perceived Self-Efficacy, Flow and Performance in a Mental Arithmetic Task. *Frontiers in psychology*, 11, 1008. <https://doi.org/10.3389/fpsyg.2020.01008>
- Pekrun, R. (2011). Emotions as Drivers of Learning and Cognitive Development. In R. A. Calvo & S. K. D'Mello (Hrsg.), *New Perspectives on Affect and Learning Technologies* (S. 23–39). Springer New York. https://doi.org/10.1007/978-1-4419-9625-1_3
- Perlwitz, P. & Stemann, J. (2022). Flow and Self-efficacy in a Serious Game for STEM Education. In H. Söbke, P. Spangenberg, P. Müller & S. Göbel (Hrsg.), *Lecture Notes in Computer Science. Serious Games* (Bd. 13476, S. 3–16). Springer International Publishing. https://doi.org/10.1007/978-3-031-15325-9_1
- Perlwitz, P., Stemann, J. & Chaari, A. (2022). Serious Games im Technikunterricht-lernwirksam oder nur Spielerei? *technik-education (tedu). Fachzeitschrift für Unterrichtspraxis und Unterrichtsforschung im allgemeinbildenden Technikunterricht*, 2(1), 3–10.
- Perttula, A., Kiili, K., Lindstedt, A. & Tuomi, P. (2017). Flow experience in game based learning: A Systematic literature review. *International Journal of Serious Games*, 4(1), 57–72. <https://doi.org/10.17083/ijsg.v4i1.151>
- Pintrich, P. R. & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. *The scholarship of teaching and learning in higher education: An evidence-based perspective*, 731–810.
- Plass, J. L., Homer, B. D. & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Potosky, D. (2002). A field study of computer efficacy beliefs as an outcome of training: the role of computer playfulness, computer knowledge, and performance during training. *Computers in Human Behavior*, 18(3), 241–255. [https://doi.org/10.1016/S0747-5632\(01\)00050-4](https://doi.org/10.1016/S0747-5632(01)00050-4)
- Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment*, 1(1).

- Prenzel, M., Häußler, P., Rost, J. & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft*, 30(2), 120–135.
- Qian, M. & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, 63, 50–58.
<https://doi.org/10.1016/j.chb.2016.05.023>
- Quandt, T., Wimmer, J. & Wolling, J. (2008). *Die Computerspieler: Studien zur Nutzung von Computergames*. VS Verlag für Sozialwissenschaften.
- Rachmatullah, A., Reichsman, F., Lord, T., Dorsey, C., Mott, B., Lester, J. & Wiebe, E. (2021). Modeling Secondary Students' Genetics Learning in a Game-Based Environment: Integrating the Expectancy-Value Theory of Achievement Motivation and Flow Theory. *Journal of Science Education and Technology*, 30(4), 1–18.
<https://doi.org/10.1007/s10956-020-09896-8>
- Rahimi, S., Shute, V. J., Fulwider, C., Bainbridge, K., Kuba, R., Yang, X., Smith, G., Baker, R. S. & D'Mello, S. K. (2022). Timing of learning supports in educational games can impact students' outcomes. *Computers & Education*, 190, 104600.
- Rasch, D., Kubinger, K. D. & Yanagida, T. (2011). *Statistics in Psychology Using R and SPSS*. Wiley.
<https://doi.org/10.1002/9781119979630>
- Ratan, R. & Ritterfeld, U. (2009). Classifying Serious Games. In U. Ritterfeld, M. J. Cody & P. Vorderer (Hrsg.), *Serious Games: Mechanisms and Effects* (S. 32–46). Routledge.
<https://doi.org/10.4324/9780203891650-10>
- Ravayse, W. S., Seugnet Blignaut, A., Leendertz, V. & Woolner, A. (2017). Success factors for serious games to enhance learning: a systematic review. *Virtual Reality*, 21(1), 31–58.
<https://doi.org/10.1007/s10055-016-0298-4>
- Reeves, B. & Read, J. L. (2009). *Total engagement: How games and virtual worlds are changing the way people work and businesses compete*. Harvard Business Press.
- Revelle, W. & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154.
- Rheinberg, F., Iser, I. & Pfauter, S. (1997). Freude am Tun und/oder zweckorientiertes Schaffen - Zur transsituativen Konsistenz und konvergenten Validität der Anreiz-Fokus-Skala. *Diagnostica*, 2, 174–191.
- Rheinberg, F. & Vollmeyer, R. (2003). Flow-Erleben in einem Computerspiel unter experimentell variierten Bedingungen. *Zeitschrift für Psychologie*, 211(4), 161–170.
<https://doi.org/10.1026//0044-3409.211.4.161>
- Rheinberg, F., Vollmeyer, R. & Engeser, S. (2019). *FKS - Flow-Kurzskala [Verfahrensdokumentation, Fragebogen und Normtabelle]*. Leibniz-Institut für Psychologie (ZPID) (Hrsg.), Open Test Archive.
<https://doi.org/10.23668/psycharchives.2667>
- Rhöneck, C. von (1986). Vorstellungen vom elektrischen Stromkreis und zu den Begriffen Strom, Spannung und Widerstand. *Naturwissenschaften im Unterricht-Physik*, 34(13), 10–14.
- Rincke, K. & Leisen, J. (2015). Sprache im Physikunterricht. In E. Kircher, R. Girwidz & P. Häußler (Hrsg.), *Physikdidaktik: Theorie und Praxis* (S. 635–655). Springer.
https://doi.org/10.1007/978-3-642-41745-0_21
- Riopel, M., Nenciovici, L., Potvin, P., Chastenay, P., Charland, P., Sarrasin, J. B. & Masson, S. (2019). Impact of serious games on science learning achievement compared with more conventional instruction: an overview and a meta-analysis. *Studies in Science Education*, 55(2), 169–214. <https://doi.org/10.1080/03057267.2019.1722420>
- Riva, E. F. M., Riva, G., Talo, C., Boffi, M., Rainisio, N., Pola, L., Diana, B., Villani, D., Argenton, L. & Inghilleri, P. (2017). Measuring dispositional flow: validity and reliability of the dispositional flow state scale 2, Italian version. *PloS one*, 12(9), e0182201.

- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational measurement: issues and practice*, 24(2), 3–13.
- Rodríguez-Sánchez, A., Salanova, M., Cifre, E. & Schaufeli, W. B. (2011). When good is good: A virtuous circle of self-efficacy and flow at work among teachers. *Revista de Psicología Social*, 26(3), 427–441.
- Rosenthal, S. B., Ratan, R., Rosenthal, S. & Ratan, R. A. (2022). Balancing Learning and Enjoyment in Serious Games: Kerbal Space Program and the Communication Mediation Model // Balancing learning and enjoyment in serious games: Kerbal Space Program and the communication mediation model. *Computers & Education*, 182(7), Artikel 104480. <https://doi.org/10.1016/j.compedu.2022.104480>
- Rossin, D., Ro, Y. K., Klein, B. D. & Guo, Y. M. (2009). The effects of flow on learning outcomes in an online information management cou. *Journal of information systems education*, 20(1), 87–98.
- Rost, D. H., Dickhäuser, O., Sparfeldt, J. R. & Schilling, S. R. (2004). Fachspezifische Selbstkonzepte und Schulleistungen im dimensionalen Vergleich: Eine versuchsplanerische Überprüfung des I/E-Modells. *Zeitschrift für Pädagogische Psychologie*, 18(1), 43–52.
- Rost, D. H. & Sparfeldt, J. R. (2002). Facetten des schulischen Selbstkonzepts. *Diagnostica*, 48(3), 130–140. <https://doi.org/10.1026//0012-1924.48.3.130>
- Rost, J. (2004). Psychometrische Modelle zur Überprüfung von Bildungsstandards anhand von Kompetenzmodellen - ZfPaed_2004_5_Rost_Psychometrische_Modelle_zur_Ueberpruefung_D_A.pdf. *Zeitschrift für Pädagogik*, 50(5), 662–678.
- Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL <http://www.jstatsoft.org/v48/i02/>
- RStudio Team. (2023). *RStudio: Integrated Development for R [Computer software]*. RStudio: Integrated Development for R
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*, 17(4), 688–690.
- Ryan, R. M., Rigby, C. S. & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30(4), 344–360.
- Sachser, N. (2004). Neugier, Spiel und Lernen: Verhaltensbiologische Anmerkungen zur Kindheit. *Zeitschrift für Pädagogik*, 50(4), 475–486.
- Salanova, M., Bakker, A. B. & Llorens, S. (2006). Flow at work: Evidence for an upward spiral of personal and organizational resources. *Journal of Happiness studies*, 7(1), 1–22.
- Satorra, A. & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Hrsg.), *Latent variables analysis: Applications for developmental research* (S. 399–419). SAGE Publications.
- Sawyer, B. (2007). Serious Games: Broadening Games Impact Beyond Entertainment. *Computer Graphics Forum*, 26(3), xviii. <https://doi.org/10.1111/j.1467-8659.2007.01044.x>
- Sawyer, B. & Rejeski, D. (2002). *Serious Games: Improving Public Policy through Game-Based Learning and Simulation*. Woodrow Wilson International Center for Scholars.
- Schecker, H. & Duit, R. (2018). Schülervorstellungen und Physiklernen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 1–21). Springer; Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57270-2_1
- Schiefele, U. & Roussakis, E. (2006). Die Bedingungen des Flow-Erlebens in einer experimentellen Spielsituation. *Zeitschrift für Psychologie / Journal of Psychology*, 214(4), 207–219. <https://doi.org/10.1026/0044-3409.214.4.207>
- Schiefele, U. & Schreyer, I. (1994). Intrinsische Lernmotivation und Lernen. Ein Überblick zu Ergebnissen der Forschung. *Zeitschrift für Pädagogische Psychologie*.

- Schiepe-Tiska, A., Heine, J.-H. & Lüdtke, O. (2016). Mehrdimensionale Bildungsziele im Mathematikunterricht und ihr Zusammenhang mit den Basisdimensionen der Unterrichtsqualität. *Unterrichtswissenschaft*, 3(44).
- Schlittgen, R. (2013). *Regressionsanalysen mit R*. Oldenbourg Wissenschaftsverlag Verlag.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological assessment*, 8(4), 350–353. <https://doi.org/10.1037/1040-3590.8.4.350>
- Schrader, C. & Bastiaens, T. J. (2012). The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. *Computers in Human Behavior*, 28(2), 648–658.
- Schunk, D. H. (1984). Sequential attributional feedback and children's achievement behaviors. *Journal of Educational Psychology*, 76(6), 1159–1169. <https://doi.org/10.1037/0022-0663.76.6.1159>
- Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, 25(1), 71–86.
- Schunk, D. H. (1991). Self-Efficacy and Academic Motivation. *Educational Psychologist*, 26(3-4), 207–231. <https://doi.org/10.1080/00461520.1991.9653133>
- Schunk, D. H. (2012). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra & J. Sweller (Hrsg.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (S. 101–123). American Psychological Association. <https://doi.org/10.1037/13273-005>
- Schunk, D. H., Hanson, A. R. & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, 79(1), 54–61. <https://doi.org/10.1037/0022-0663.79.1.54>
- Schunk, D. H. & Mullen, C. A. (2012). Self-Efficacy as an Engaged Learner. In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of Research on Student Engagement* (S. 219–235). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_10
- Schwarzer, M. & Tschauko, O. (2010). *Experimente bei Physikscharbeiten und Tests* (Fonds für Unterrichts- und Schulentwicklung S5 „Entdecken, Forschen und Experimentieren“).
- Schwarzer, R. (1998). Self-Science: Das Trainingsprogramm zur Selbstführung von Lehrern. *Unterrichtswissenschaft*, 26. <https://doi.org/10.25656/01:7771> (Unterrichtswissenschaft 26 (1998) 2, S. 158-172).
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (Bd. 44, S. 28–53). Beltz. <https://doi.org/10.25656/01:3930>
- Sedig, K. (2007). Toward operationalization of ‘flow’ in mathematics learnware. *Computers in Human Behavior*, 23(4), 2064–2092. <https://doi.org/10.1016/j.chb.2006.11.001>
- Seel, N. M. (2000). *Psychologie des Lernens: Lehrbuch für Pädagogen und Psychologen; mit 17 Tabellen und zahlreichen Übungsaufgaben*. Reinhardt.
- Seelhammer, C. & Niegemann, H. M. (2009). Playing games to learn—Does it actually work? *Proc. ICCE2009*, 675–681.
- Seidel, T., Prenzel, M., Rimmel, R., Dalehefte, I. M., Herweg, C., Kobarg, M. & Schwindt, K. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, 52(6), 799–821.
- Seligman, M. E., Nolen-Hoeksema, S., Thornton, N. & Thornton, K. M. (1990). Explanatory Style as a Mechanism of Disappointing Athletic Performance. *Psychological science*, 1(2), 143–146. <https://doi.org/10.1111/j.1467-9280.1990.tb00084.x>
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shavelson, R. J. & Bolus, R. (1982). Self concept: The interplay of theory and methods. *Journal of Educational Psychology*, 74(1), 3–17. <https://doi.org/10.1037/0022-0663.74.1.3>

- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of educational research*, 46(3), 407–441.
- Shernoff, D. J., Csikszentmihalyi, M., Schneider, B. & Shernoff, D. J. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18(2), 158–176.
- Shernoff, D. J., Hamari, J. & Rowe, E. (2014). Measuring Flow in Educational Games and Gamified Learning Environments. In Herrington, Jan, Ed., Viteli, Jarmo, Ed. & Leikmaa, Marianna, Ed. (Hrsg.), *EdMedia 2014: World Conference on Educational Media & Technology (Tampere, Finland, June 23-26, 2014)* (S. 2276–2281). Association for the Advancement of Computing in Education.
- Shute, V. J., D'Mello, S. K., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., Ventura, M. & Almeda, V. (2015). Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. *Computers & Education*, 86, 224–235.
- Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel psychology*, 64(2), 489–528. <https://doi.org/10.1111/j.1744-6570.2011.01190.x>
- Sitzmann, T. & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychological bulletin*, 137(3), 421–442. <https://doi.org/10.1037/a0022777>
- Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology*, 89(1), 71.
- Skaalvik, E. M. & Rankin, R. J. (1995). A Test of the Internal/External Frame of Reference Model at Different Levels of Math and Verbal Self-Perception. *American Educational Research Journal*, 32(1), 161–184. <https://doi.org/10.3102/00028312032001161>
- Skadberg, Y. X. & Kimmel, J. R. (2004). Visitors' flow experience while browsing a Web site: its measurement, contributing factors and consequences. *Computers in Human Behavior*, 20(3), 403–422. [https://doi.org/10.1016/S0747-5632\(03\)00050-5](https://doi.org/10.1016/S0747-5632(03)00050-5)
- Sokoloff, D. R. (1996). Teaching electric circuit concepts using microcomputer-based current/voltage probes. In R. F. Tinker (Hrsg.), *Microcomputer-Based Labs: Educational Research and Standards* (S. 129–146). Springer.
- Solga, H. & Pfahl, L. (2009). Doing Gender im Technisch-Naturwissenschaftlichen Bereich. In Joachim Milberg (Hrsg.), *Förderung des Nachwuchses in Technik und Naturwissenschaft: Beiträge zu den Zentralen Handlungsfeldern* (S. 155–218). Springer.
- Spybrook, J. & Raudenbush, S. W. (2009). An Examination of the Precision and Technical Accuracy of the First Wave of Group-Randomized Trials Funded by the Institute of Education Sciences. *Educational Evaluation and Policy Analysis*, 31(3), 298–318. <https://doi.org/10.3102/0162373709339524>
- Squire, K. D., Barnett, M., Grant, J. M. & Higginbotham, T. (2004). Electromagnetism supercharged! Learning physics with digital simulation games. *ICLS '04: Proceedings of the 6th international conference on Learning sciences*, 513–520.
- Squire, K. D. & Jenkins, H. (2003). Harnessing the power of games in education. *Insight*, 3(1), 5–33.
- Starks, K. (2014). Cognitive behavioral game design: a unified model for designing serious games. *Frontiers in psychology*, 5, 28.
- Stege, L., van Lankveld, G. & Spronck, P. (2012). Teaching high school physics with a serious game. *International Journal of Computer Science in Sports*, 11(1), 123–134.

- Strömmer, T. & Winkelmann, J. (2020). Charakteristische Merkmale von Physikunterricht: Wirkung auf (Un-) Beliebtheit, Interesse und Schwierigkeit. *PhyDid B-Didaktik der Physik-Beiträge zur DPG-Frühjahrstagung*.
- Strzebkowski, R. & Kleeberg, N. (2002). Interaktivität und Präsentation als Komponenten multimedialer Lernanwendungen. In L. J. Issing (Hrsg.), *Information und Lernen mit Multimedia und Internet* (Bd. 3, S. 229–245). Verl. Internat. Psychoanalyse.
- Susi, T., Johannesson, M. & Backlund, P. (2007). *Technical Report HS- IKI -TR-07-001: Serious games: An overview*. School of Humanities and Informatics University of Skövde, Sweden.
- Sweetser, P. & Wyeth, P. (2005). GameFlow:: A Model for Evaluating Player Enjoyment in Games. *ACM Computers in Entertainment*, 3(3), 1–24.
- Sweller, J. (2005). Implications of Cognitive Load Theory for Multimedia Learning. In R. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 19–30). Cambridge University Press. <https://doi.org/10.1017/CB09780511816819.003>
- Szogs, M., Korneck, F., Krüger, M., Oettinghaus, L. & Kunter, M. (2016). Kognitive Aktivierung in standardisierten Unterrichtsminiaturen. *Authentizität und Lernen–das Fach in der Fachdidaktik*, 605–610.
- Tabachnick, B. G. & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Thomson/Brooks/Cole Belmont, CA.
- Tausch, A. & Peifer, C. (2019). Auswirkungen von Autonomie auf Flow, Motivation und Leistung: Eine Studie im Schaltanlagenbau. *Wirtschaftspsychologie*, 21(4), 83–100.
- Toh, W. & Kirschner, D. (2020). Self-directed learning in video games, affordances and pedagogical implications for teaching and learning. *Computers & Education*, 154, 103912.
- Toppo, G. (2021). *Oregon Trail Began 50 Years Ago in Minneapolis Classroom*. <https://www.the74million.org/article/oregon-trail-at-50-how-three-teachers-created-the-computer-game-that-inspired-and-diverted-generations-of-students/>
- Tsai, M.-J., Huang, L.-J., Hou, H.-T., Hsu, C.-Y. & Chiou, G.-L. (2016). Visual behavior, flow and achievement in game-based learning. *Computers & Education*, 98, 115–129.
- Ullah, M., Amin, S., Munsif, M., Safaev, U., Khan, H., Khan, S. & Ullah, H. (2022). Serious Games in Science Education. A Systematic Literature Review. *Virtual Reality & Intelligent Hardware*, 4(3), 189–209. <https://doi.org/10.1016/j.vrih.2022.02.001>
- Urban, D. & Mayerl, J. (2018). *Angewandte Regressionsanalyse: Theorie, Technik und Praxis*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-01915-0>
- Urbani, S., Smith, M. R., Maddux, C. D., Smaby, M. H., Torres-Rivera, E. & Crews, J. (2002). Skills-Based Training and Counseling Self-Efficacy. *Counselor Education and Supervision*, 42(2), 92–106. <https://doi.org/10.1002/j.1556-6978.2002.tb01802.x>
- Urban-Woldron, H. & Hopf, M. (2012). Entwicklung eines Testinstruments zum Verständnis in der Elektrizitätslehre Development of a diagnostic instrument for testing student understanding of basic electricity concepts. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 203–229.
- Usher, E. L. & Pajares, F. (2008). Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions. *Review of educational research*, 78(4), 751–796. <https://doi.org/10.3102/0034654308321456>
- Utts, J. M. (1982). The rainbow test for lack of fit in regression. *Communications in Statistics-Theory and Methods*, 11(24), 1801–1815.
- Veermand, K. & Jaakkola, T. (2019). Pedagogy in Educational Simulations and Games. In Y. Cai, W. R. van Joolingen & Z. Walker (Hrsg.), *Gaming Media and Social Effects. VR, Simulations and Serious Games for Education* (S. 5–14). Springer Singapore. https://doi.org/10.1007/978-981-13-2844-2_2
- Viering, T., Fischer, H. E. & Neumann, K. (2010). Die Entwicklung physikalischer Kompetenz in der Sekundarstufe I. Projekt Physikalische Kompetenz. In E. Klieme, D. Leutner & M.

- Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (S. 92–103). Beltz.
- Virvou, M., Katsionis, G. & Manos, K. (2005). Combining software games with education: Evaluation of its Educational Effectiveness. *Journal of Educational Technology & Society*, 8(2), 54–65.
- Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A. & Muse, K. (2006). Computer Gaming and Interactive Simulations for Learning: A Meta-Analysis. *Journal of Educational Computing Research*, 34(3), 229–243.
- Vollmeyer, R. & Imhof, M. (2007). Are there gender differences in computer performance? If so, can motivation explain them? *Zeitschrift für Pädagogische Psychologie*, 21(3/4), 251–261.
- Vorderer, P., Klimmt, C. & Ritterfeld, U. (2004). Enjoyment: At the Heart of Media Entertainment. *Communication Theory*, 14(4), 388–408. <https://doi.org/10.1111/j.1468-2885.2004.tb00321.x>
- Walpuski, M. & Ropohl, M. (2014). Statistische Verfahren für die Analyse des Einflusses von Aufgabenmerkmalen auf die Schwierigkeit. *Methoden in der naturwissenschaftsdidaktischen Forschung*, 385–398.
- Wang, C.-C. & Hsu, M.-C. (2014). An exploratory study using inexpensive electroencephalography (EEG) to understand flow experience in computer-based instruction. *Information & Management*, 51(7), 912–923.
- Wang, C.-H., Wu, K.-C. & Tsau, S.-Y. (2018). Flow Learning Experience: Applying Marketing Theory to Serious Game Design. *Journal of Educational Computing Research*, 57(2), 417–447. <https://doi.org/10.1177/0735633117752454>
- Wang, J., Stebbins, A. & Ferdig, R. E. (2022). Examining the effects of students' self-efficacy and prior knowledge on learning and visual behavior in a physics game. *Computers & Education*, 178, Artikel 104405. <https://doi.org/10.1016/j.compedu.2021.104405>
- Wastiau, P., Kearney, C. & Van den Berghe, W. (2009). *How are digital games used in schools?* European Schoolnet.
- Webb-Williams, J. (2018). Science self-efficacy in the primary classroom: Using mixed methods to investigate sources of self-efficacy. *Research in Science Education*, 48(5), 939–961.
- Weimar, D. (2005). *Streß und Flow-Erleben: eine empirische Untersuchung zur Bedeutung von Kognitionen, Emotionen und Motivation bei Lehramtsstudierenden, Referendaren und Lehrern*. Logos.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. 08893724.
- Welberg, J., Laumann, D. & Heinicke, S. (2021). „Und für wen ist dieser Kontext?“-Studien zu Kontexten und Interessen im Physikunterricht unter Beachtung von Gender und Selbstkonzept. *PhyDid B-Didaktik der Physik-Beiträge zur DPG-Frühjahrstagung*.
- Wentzel, K. R. & Miele, D. B. (2009). *Handbook of motivation at school*. Routledge.
- Wenzel, M. (2018). *Computereinsatz in Schule und Schülerlabor: Einstellung von Physiklehrkräften zu Neuen Medien. Studien zum Physik- und Chemielernen*, Bd. 251. Logos Verlag. <https://doi.org/10.30819/4659>
- Werbach, K., Hunter, D. & Dixon, W. (2012). *For the win: How game thinking can revolutionize your business*. Wharton digital press Philadelphia.
- Werner, C. S. (2015). *Strukturgleichungsmodelle mit R und lavaan analysieren: Kurzeinführung*.
- Weßnigk, S. (2012). *Kooperatives Arbeiten an industrienahen außerschulischen Lernorten* [Dissertation Naturwissenschaften, Christian-Albrechts-Universität Kiel, Kiel]. RIS. https://macau.uni-kiel.de/receive/diss_mods_00010788.
- Wetzel, I. (2002). Teaching Computer Skills: A Gendered Approach. In C. Floyd (Hrsg.), *Feminist Challenges in the Information Age* (S. 223–239). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-94954-7_17

- Whitley, B. E. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in Human Behavior*, 13(1), 1–22. [https://doi.org/10.1016/S0747-5632\(96\)00026-X](https://doi.org/10.1016/S0747-5632(96)00026-X)
- Whitson, C. & Consoli, J. (2009). Flow theory and student engagement. *Journal of Cross-Disciplinary Perspectives in Education*, 2(1), 40–49.
- Whitton, N. (2007). Motivation and computer game based learning. In *Proceedings ascilite Singapore*.
- Wilhelm, T. & Hopf, M. (2018). Schülervorstellungen zum elektrischen Stromkreis: Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 115–138). Springer; Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57270-2_6
- Wilhelm, T. & Vairo Nunes, R. M. (2019). Vergleichende Schulbuchanalyse zur Einführung in die E-Lehre. In S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Wien 2019* (S. 578–581). Universität Duisburg-Essen.
- Wilkinson, P. (2016). A Brief History of Serious Games. In R. Dörner, S. Göbel, M. Kickmeier, M. Masuch & K. Zweig (Hrsg.), *Entertainment Computing and Serious Games: International GI-Dagstuhl Seminar 15283, Dagstuhl Castle, Germany, July 5-10, 2015, Revised Selected Papers* (S. 17–41). Springer.
- Wirth, J. & Leutner, D. (2005). Selbstregulation beim Lernen in interaktiven Lernumgebungen. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 172–184). Hogrefe.
- Wolff, F., Helm, F., Zimmermann, F., Nagy, G. & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology*, 110(7), 1005–1025. <https://doi.org/10.1037/edu0000248>
- Woodworth, R. S. (1918). *Dynamic psychology*. Columbia University Press.
- Wouters, P., van der Spek, E. D. & van Oostendorp, H. (2009). Current practices in serious game research: A review from a learning outcomes perspective. *Games-based learning advancements for multi-sensory human computer interfaces: techniques and effective practices*, 232–250.
- Wouters, P., van Nimwegen, C., van Oostendorp, H. & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- Wylie, R. C. (1968). The present status of self theory. In E. F. Borgatta & W. W. Lambert (Hrsg.), *Handbook of personality theory* (S. 728–787). Rand McNally.
- Yang, J. C., Quadir, B. & Chen, N.-S. (2016). Effects of the Badge Mechanism on Self-Efficacy and Learning Performance in a Game-Based English Learning Environment. *Journal of Educational Computing Research*, 54(3), 371–394. <https://doi.org/10.1177/0735633115620433>
- Yang, X., Zhang, M., Kong, L., Wang, Q. & Hong, J.-C. (2021). The effects of scientific self-efficacy and cognitive anxiety on science engagement with the “Question-Observation-Doing-Explanation” model during school disruption in COVID-19 pandemic. *Journal of Science Education and Technology*, 30(3), 380–393. <https://doi.org/10.1007/s10956-020-09877-x>
- Yang, Y. & Xia, Y. (2019). Categorical Omega With Small Sample Sizes via Bayesian Estimation: An Alternative to Frequentist Estimators. *Educational and Psychological Measurement*, 79(1), 19–39. <https://doi.org/10.1177/0013164417752008>
- Yee, N. & Bailenson, J. (2007). The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human communication research*, 33(3), 271–290. <https://doi.org/10.1111/j.1468-2958.2007.00299.x>

- Zhonggen, Y. (2019). A Meta-Analysis of Use of Serious Games in Education over a Decade. *International Journal of Computer Games Technology*(17), 1–8.
<https://doi.org/10.1155/2019/4797032>
- Zimmerman, B. J. (1995). Self-efficacy and educational development. *Self-efficacy in changing societies*, 1(1), 202–231.
- Zimmerman, B. J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25(1), 82–91. <https://doi.org/10.1006/ceps.1999.1016>
- Zubair, A. & Kamal, A. (2015). Authentic leadership and creativity: mediating role of work-related flow and psychological capital. *Journal of Behavioural Sciences*, 25(1).
- Zulkosky, K. (2009). Self-Efficacy: A Concept Analysis. *Nursing Forum*, 44(2), 93–102.
<https://doi.org/10.1111/j.1744-6198.2009.00132.x>
- Zumbach, J., Rammerstorfer, L. & Deibl, I. (2020). Cognitive and metacognitive support in learning with a serious game about demographic change. *Computers in Human Behavior*, 103, 120–129. <https://doi.org/10.1016/j.chb.2019.09.026>
- Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9), 25–32.

8 Anhang

A) Einordnung der Lerninhalte des Spiels

Kapitel 1 - Elektrizität in unserem Alltag

Lernziele: Die Schülerinnen und Schüler (SuS) können...

Bildungsplan:

- (1) grundlegende Bauteile eines elektrischen Stromkreises benennen und ihre Funktion beschreiben (unter anderem Schaltsymbole)
- (5) den Aufbau eines Stromkreises unter Vorgabe einer Schaltskizze durchführen sowie Stromkreise in Form von Schaltskizzen darstellen. (MKJS, 2015b)

Eigene Formulierung:

- die Elektrizität in ihrem Alltag erkennen.
- den Begriff "Elektrizität" einordnen.
- Polstellen an elektrischen Geräten erkennen.
- Batterie und Netzgerät als Quelle der Elektrizität erkennen.
- die Bauteile des einfachen el. Stromkreises und ihre Symbole nennen (Nutzer, Leiter, verschiedene Elektrizitätsquellen),
- den Kurzschluss in einem elektrischen Stromkreis erkennen,
- den Elektrizitätsantrieb als Antrieb für die Elektrizität erkennen,

Aufgaben Kinderzimmer

- Welche Geräte im Haus sind vom Stromausfall betroffen?
- Welche Teile benötigt man für einen einfachen elektrischen Stromkreis?
- Welche Symbole stellen ___ dar?
- Welche Voraussetzungen müssen erfüllt sein, damit eine Lampe in einem einfachen Stromkreis leuchten kann?
- Bei welchen Schaltungen gibt es einen Kurzschluss?
- Was passiert im Stromkreis bei einem Kurzschluss?

Texte im Wohnzimmer

- Der elektrische Stromkreis
- Der Kurzschluss
- Der Elektrizitätsantrieb
- Symbole in Schaltungen

Versuche auf dem Dachboden

- Einen einfachen el. Stromkreis bauen
- Einen Kurzschluss erzeugen
- Zwei Lampen zum Leuchten bringen

Kapitel 2 - Die elektrische Stromstärke

Lernziele: Die SuS können...

Bildungsplan:

- (3) qualitativ beschreiben, dass elektrische Ströme einen Antrieb beziehungsweise eine Ursache benötigen und durch Widerstände in ihrer Stärke beeinflusst werden (Stromstärke, Potential, Spannung, Widerstand, Ladung).
- (5) den Aufbau eines Stromkreises unter Vorgabe einer Schaltskizze durchführen sowie Stromkreise in Form von Schaltskizzen darstellen.
- (6) Stromstärke und Spannung messen
- (9) physikalische Angaben auf Alltagsgeräten beschreiben (Spannung, Stromstärke, Leistung). (MKJS, 2015b)

Eigene Formulierung:

- die Stromstärke im Stromkreis ermitteln.
- die Stärke des elektrischen Stroms messen.
- die Stärke in einer Reihenschaltung messen.
- erkennen, dass die Stromstärke in einer Reihenschaltung in allen Bauteilen gleich ist.
- die das Formelzeichen I und die Einheit Ampère (1 A) der elektrischen Stromstärke nennen.
- den Zusammenhang, je größer der Widerstand eines Bauteils ist, desto kleiner ist die Stromstärke, nachvollziehen.

Aufgaben Kinderzimmer

- Was trifft auf die physikalische Größe elektrische Stromstärke zu?
- In welcher Schaltung werden die Lampen am hellsten leuchten?
- Bei welchen Schaltungen ist das Amperemeter so angeschlossen, dass wir die Stromstärke im Stromkreis auslesen können?
- In welchem Stromkreis wird die Stromstärke am höchsten sein?

Texte im Wohnzimmer

- Die elektrische Stromstärke
- Messen des elektrischen Stroms - Das Amperemeter
- Die Reihenschaltung

Versuche auf dem Dachboden

- Reihenschaltung bauen
- Stromstärke messen
- Lampen kombinieren, bis sie eine vorgegebene Stromstärke erreichen
- Lampen heller leuchten lassen

Kapitel 3 - Die elektrische Spannung

Lernziele: Die SuS können...

Bildungsplan:

- (3) qualitativ beschreiben, dass elektrische Ströme einen Antrieb beziehungsweise eine Ursache benötigen und durch Widerstände in ihrer Stärke beeinflusst werden (Stromstärke, Potential, Spannung, Widerstand, Ladung).
- (5) den Aufbau eines Stromkreises unter Vorgabe einer Schaltskizze durchführen sowie Stromkreise in Form von Schaltskizzen darstellen.
- (6) Stromstärke und Spannung messen.
- (9) physikalische Angaben auf Alltagsgeräten beschreiben (Spannung, Stromstärke, Leistung). (MKJS, 2015b)

Eigene Formulierung:

- die elektrische Spannung als Antrieb für den elektrischen Strom erklären.
- erkennen, dass die Spannung U einer Elektrischen Energiequelle gleich dem Potentialunterschied von Plus- und Minuspol ist.
- die Einheit von Spannung und Potenzial als das Volt (1 V) benennen.
- erkennen, dass In Parallelschaltungen die Spannung bei allen parallelen Bauteilen gleich groß ist.
- die Spannung gemessen.
- die Proportionalität der Spannung zum Widerstand des Nutzers erkennen.

Aufgaben Kinderzimmer

- Bei welchen Schaltungen wirst du die gleiche Spannung bei allen Lampen messen können? Bei welcher Schaltung ist das Voltmeter richtig eingebaut, um die Spannung an der markierten Lampe zu messen?
- Wie hoch ist das Potential (in Volt) an der angegebenen Stelle?
- Welche Aussagen über die el. Spannung treffen zu?
- Kombinationen von Batterien – Spannung ermitteln

Texte im Wohnzimmer

- Das Potenzial
- Die elektrische Spannung
- Messen der elektrischen Spannung bzw. des Potentialunterschiedes
- Parallelschaltungen

Versuche auf dem Dachboden

- Parallelschaltung
- Spannung messen
- Vergleich von Parallel- und Reihenschaltung

Differenzierung: Kapitel 4 - Schaltungen verstehen und planen

Lernziele: Die SuS können...

Bildungsplan:

- (5) den Aufbau eines Stromkreises unter Vorgabe einer Schaltskizze durchführen sowie Stromkreise in Form von Schaltskizzen darstellen.
- (6) Stromstärke und Spannung messen.
- (7) in einfachen Reihenschaltungen und Parallelschaltungen Gesetzmäßigkeiten für die Stromstärke und die Spannung beschreiben (Maschenregel, Knotenregel). (MKJS, 2015b)

Eigene Formulierung:

- erkennen, dass die Summe der Spannungen an den einzelnen Geräten gleich der Spannung der Elektrizitätsquelle ist:
 $U_1+U_2+U_3+\dots=U_Q$.
- erkennen, dass die Summe der Stromstärken in den einzelnen Geräten gleich der Gesamtstromstärke ist:
 $I_1+I_2+I_3+\dots=I_{\text{Gesamt}}$.

Aufgaben im Kinderzimmer

- Bei welcher Anordnung leuchten alle Lampen gleich hell?
- Bei welcher Anordnung leuchten die Lampen am hellsten?
- Was wird das Amperemeter anzeigen?
- Was wird das Voltmeter anzeigen?

Wohnzimmer Texte

- Die Maschenregel
- Die Knotenregel

Versuche auf dem Dachboden

- Kombination aus Reihen- und Parallelschaltung
- Versuch zu Reihenschaltung und Spannung
- Versuch zu Parallelschaltung und Stromstärke
- Versuch zur Schaltung von Batterien

B) Bewertungsbogen

Dein Code				

Du kannst, musst aber nichts ausfüllen. Jedoch kann jede Antwort helfen, Spiele für den Unterricht besser zu machen.

- Diese Fragen/Begriffe/Formulierungen habe ich nicht verstanden:
- Hier finde ich die Schrift unleserlich:
- Das gefällt mir gar nicht:
- Das finde ich gut:
- Diese Aufgaben fand ich zu schwer (im Format: Kapitel/Raum/Aufgabennummer):

Im späteren Verlauf des Spiels wird immer wieder ein Fenster auftauchen, bei dem du nach deinem aktuellen Spiel-Gefühl gefragt wirst.

- Bei dieser Frage wusste ich nicht, was ich antworten soll:
- Diese Frage habe ich nicht verstanden:
- Was ist dir sonst aufgefallen, was für dich wichtig ist?

Testfall	OK	Nicht OK (kurze Beschreibung)
Das Spiel ließ sich unkompliziert starten.		
Die Grafik lief flüssig/kein ruckeln/kein flackern.		
Die Schrift war bei Vollbild gut lesbar.		
Die Spieloberfläche ist übersichtlich und ich konnte mich schnell zurechtfinden		
Ich wusste, was ich in dem Spiel zu tun habe.		
Die Fragen zwischendurch haben mich nicht gestört.		
Ich habe das Gefühl, durch das Spiel etwas über die Elektrizität gelernt zu haben.		
Das Spiel war abwechslungsreich.		
<p>Diese Dinge sind mir positiv aufgefallen:</p> <p>Diese Dinge sind mir negativ aufgefallen:</p>		
Das wollte ich noch loswerden:		

C) Soci Survey Fragebogenseiten



0% ausgefüllt

Liebe Schülerin, lieber Schüler,

Ich bin eine Doktorandin an der pädagogischen Hochschule in Freiburg und möchte gerne in deiner Klasse untersuchen, wie gut ihr mit digitalen Lernspielen im Physikunterricht lernt. Für die Untersuchung möchte ich dich bitten, einen Fragebogen vor und nach dem Lernspiel auszufüllen.

Deine Eltern haben bereits zugestimmt, dass du die Fragen beantworten darfst. Trotzdem ist deine Teilnahme freiwillig. Es entstehen dir keine Nachteile, wenn du nicht teilnehmen möchtest.

Deinen Antworten werden anonym behandelt. Das heißt, es weiß nach der Befragung niemand genau, wer von euch wie geantwortet hat.

Du kannst daher die Fragen in dem Fragebogen ganz ehrlich beantworten.

Wenn du Fragen zu dem Fragebogen hast, kannst du dich an mich oder deine Lehrkraft wenden

Ich hoffe, dass du mich bei meiner Befragung unterstützt, indem du die Fragen beantwortest.

Viele Grüße

Phoebe Perlwitz

Weiter



5% ausgefüllt

Wie lautet dein 5-stelliger Code?

Mein Code lautet:

Wie alt bist du?

Wie alt bist du?

Bist du...

männlich

weiblich

diverse

Ist Deutsch deine Muttersprache?

ja

nein

Weiter

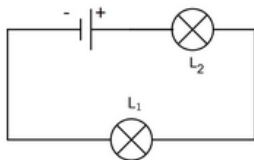
[Phoebe Perlwitz](#), Pädagogische Hochschule Freiburg – 2021

Bitte beantworte folgende Fragen:



- | | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Ich traue mir einiges zu in Physik. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Das Fach Physik liegt mir nicht besonders | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Bei manchen Sachen weiß ich schon im Voraus, dass ich sie nie verstehe, in Physik | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Ich glaube, dass mich meine Mitschüler für wirklich gut halten, in Physik | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Kein Mensch kann alles – ich habe einfach keine Begabung für Physik | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Wenn ich eine neue Aufgabe bearbeiten soll, bin ich sicher, dass ich sie schaffen werde in Physik | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Wenn eine Aufgabe kompliziert und schwierig erscheint, spornt mich das erst recht an, die Aufgabe zu lösen, in Physik | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Weiter



Betrachte die Abbildung.

Wie hell werden die Glühlampen L1 und L2 leuchten?

- Beide Glühlampen leuchten. L2 leuchtet heller als L1.
- Beide Glühlampen leuchten gleich hell.
- L2 leuchtet. L1 leuchtet nicht.
- L1 leuchtet. L2 leuchtet nicht.

Wie begründest du deine Entscheidung?

- Der elektrische Strom ist überall im Stromkreis gleich.
- Der Strom wird gleichmäßig auf beide Glühlampen aufgeteilt.
- L2 ist näher bei der Batterie. Daher bekommt sie mehr Strom.
- L2 verbraucht einen Teil des Stroms. Es ist daher weniger Strom für L1 übrig.

D) Items des Fachwissenstests

Kapitel 1: Einfacher elektrischer Stromkreis, Glühlampen, Pole, Elektrizitätsantrieb, Kurzschluss, Schaltzeichen					
Item	Frage	KB	Quelle	1-stufig/ 2-stufig	SFV
A1	Welche Bauteile benötigt man für den Bau eines einfachen elektrischen Stromkreises?	1	P	1	G
A2	Was passiert in einem einfachen elektrischen Stromkreis bei einem Kurzschluss?		P	1	KS
A3	In welchen Stromkreisen leuchten die Glühlampen ?		P	1	SB
B1	Welche Aufgabe hat die Batterie in einem elektrischen Stromkreis?		P	1	G
B3	In welchen Stromkreisen kommt es zu einem Kurzschluss?		P	1	G
Kapitel 2: Elektrische Stromstärke, Reihenschaltungen, messen der elektrischen Stromstärke					
Item	Frage	KB	Quelle	1-stufig/ 2-stufig	SFV
A4	In einem Stromkreis 1 mit einer Glühlampe 1 beträgt die Stromstärke $I = 1 \text{ A}$. Der Stromkreis wird nun durch eine weitere Glühlampe 2 ergänzt. Welchen Einfluss hat das auf die Stromstärke durch die Glühlampe 1?	2	H13	1	SV, SA, IR
A5	Wie hell werden die Glühlampen leuchten?	2	H21	2	SV, SA, IR
A6	Im Stromkreis 1 befindet sich eine Glühlampe. Sie wird durch eine zweite Glühlampe ergänzt. Was geschieht mit der Stromstärke im Stromkreis 2?	2	H6	2	SV, SA
A7	Im Stromkreis 1 leuchtet die Glühlampe und das Amperemeter A1 zeigt eine Stromstärke von $0,2 \text{ A}$ an. Nun wird ein zweites Amperemeter A2 auf der anderen Seite des Stromkreises eingebaut (Stromkreis 2). Was zeigt das Amperemeter A2 in Stromkreis 2 an?	2	H28	2	SV, SA
A8	Im Stromkreis 1 ist eine Glühlampe an eine Batterie angeschlossen. Was passiert mit der Helligkeit der Glühlampe, wenn eine zweite Batterie in Reihe geschaltet wird?	2	P3	2	SA, BS
A9	Dem Stromkreis 1 wird eine zweite, gleiche Glühlampe so hinzugefügt, wie es in Stromkreis 2 dargestellt ist. Vergleiche die Helligkeit der Glühlampe L_1 in Stromkreis 1 und 2.	2	S2	2	SA, BS
B2	Welche Aussage trifft auf die elektrische Stromstärke zu?	1	P7	1	G

B4	Dem Stromkreis 1 wird eine zweite, gleiche Glühlampe so hinzugefügt, wie es in Stromkreis 2 dargestellt ist. Vergleiche die Helligkeit der Glühlampe L_1 in den beiden Stromkreisen.	2	S1	1	SA, BS
B5	Wie hell werden die Glühlampen leuchten?	2	H11	2	SA, BS
B6	Der Stromkreis 1 besteht aus einer Batterie, einer Glühlampe und einem Amperemeter. Das Amperemeter zeigt die Stromstärke an. Was passiert mit der Anzeige des Amperemeters, wenn eine zweite, gleiche Glühlampe hinzugefügt wird (Stromkreis 2)?	2	H23	2	SA, BS
B7	Im Stromkreis ist eine Batterie mit einer Glühlampe verbunden. Was kannst du über die Anzeigen der beiden Amperemeter sagen?	2	H4	2	SA, BS

Kapitel 3: Elektrische Spannung, Parallelschaltungen

Item	Frage	KB	Quelle	1-stufig/ 2-stufig	SFV
A10	Vergleiche die Anzeigen der Voltmeter V_1 und V_2 :	2	P11	2	PS, UI
A11	Vergleiche die Helligkeiten der Glühlampen L_1 , L_2 und L_3 in den beiden Stromkreisen. Welche Glühlampe oder welche Glühlampen leuchten am hellsten?	2	H24	1	PS, UI
A12	Die Glühlampe in der Abbildung leuchtet. Was kannst du über die Spannung zwischen den Punkten A und B sagen?	2	P2	2	KS, UI
A13	Welche Glühlampen in der Abbildung sind parallel geschaltet?	2	H14	1	PS
B8	Vergleiche die Spannung, die an den Glühlampen L_1 und L_2 anliegt:	2	P2	2	KS, UI
B9	Dem Stromkreis 1 wird eine zweite, identische Glühlampe so hinzugefügt, wie es in Stromkreis 2 dargestellt ist. Vergleiche die Anzeige des Voltmeters V in Stromkreis 1 und 2.	2	P10	2	KS, UI
B10	In welchem Stromkreis oder in welchen Stromkreisen sind die Glühlampen L_1 und L_2 parallel zur Batterie geschaltet?	2	H30	1	PS
B11	In welchen Stromkreisen sind die beiden Glühlampen L_1 und L_2 parallel zur Batterie geschaltet?	2	H20	1	PS

B12	Im Stromkreis 1 ist eine Glühlampe an eine Batterie mit 6 V angeschlossen. In Stromkreis 2 kommt eine zweite Batterie mit 6 V hinzu. Wie viel Volt liegen nun insgesamt an der Glühlampe an?	2	P12	2	UI
B13	Welche Spannung liegt zwischen den Punkten 1 und 2 an?	2	P1	2	PS

Anmerkungen. KB = Kompetenzbereich 1: *Wissen wiedergeben*, 2: *Wissen anwenden*, 3: *Wissen transferieren und verknüpfen*. Codierung der Quelle: H: Urban-Woldron & Hopf-Test mit Item Nr., S: Sokoloff-Test, P: eigenes Item. Zahl = Item Nr.; SVF = Schüler:innenfehlvorstellungen; G: Notwendigkeit des geschlossenen Stromkreises nicht bewusst, SV: Stromverbrauchsvorstellung, BS: Konstantstromquelle, SA: Sequenzielles Denken, UI: Mangelnde Unterscheidung zwischen Spannung und Stromstärke, PS: Mangelnde Unterscheidung zwischen Reihen- und Parallelschaltung, SB: Umsetzung Schaltbild \leftrightarrow realer Stromkreis, KS: Konzeptionelle Schwierigkeiten mit der Spannung.

E) Ergebnisse der Pilotierung des Fachwissenstests

Pretest									
EX		KG		EX		KG			
Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
AV1	.00	.00	.44	.50	BV1	.24	.44	.12	.33
AV2	.38	.51	.10	.30	BV2	.32	.48	.33	.50
AV3	.00	.00	.09	.30	BV3	.07	.26	.08	.29
AV4	.25	.44	.48	.51	BV4	.38	.51	.00	.00
AV5	.33	.48	.50	.51	BV5	.39	.49	.10	.31
AV6	.23	.42	.19	.40	BV6	.10	.30	.09	.30
AV7	.23	.43	.16	.37	BV7	.53	.51	.16	.37
AV8	.44	.50	.11	.32	BV8	.40	.50	.00	.00
AV9	.60	.55	.44	.53	BV9	.00	.00	.13	.34
AV10	.28	.45	.06	.25	BV10	.14	.36	.04	.19
AV11	.34	.48	.33	.48	BV11	.04	.21	.08	.27
AV12	.32	.47	.00	.00	BV12	.00	.00	.10	.32
AV13	.32	.48	.06	.24	BV13	.00	.00	.00	.00
Gesamt	.29	.11	.23	.12		.20	.16	.09	.06
Posttest									
EX		KG		EX		KG			
Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
AN1	.38	.51	.30	.47	BN1	.42	.51	.17	.38
AN2	.10	.31	.19	.40	BN2	.31	.48	.40	.52
AN3	.75	.44	.05	.23	BN3	.15	.38	.08	.29
AN4	.42	.50	.24	.43	BN4	.24	.44	.17	.38
AN5	.50	.51	.46	.51	BN5	.52	.50	.37	.49
AN6	.45	.50	.16	.37	BN6	.23	.42	.11	.31
AN7	.55	.51	.16	.37	BN7	.58	.51	.35	.49
AN8	.49	.50	.16	.37	BN8	.33	.48	.11	.33
AN9	.53	.51	.40	.50	BN9	.38	.50	.42	.51
AN10	.44	.51	.04	.19	BN10	.09	.29	.11	.32
AN11	.20	.41	.41	.50	BN11	.27	.45	.32	.23
AN12	.42	.50	.33	.48	BN12	.05	.22	.05	.48
AN13	.13	.34	.33	.48	BN13	.20	.41	.10	.31
Gesamt	.41	.13	.25	.11	Gesamt	.29	.12	.21	.08

Anmerkungen. EX = Experimenttaggruppe. KG = Kontrollgruppe, M = Mittelwert, SD = Standardabweichung.
 AV: Fachwissen(t₁) der Gruppe A, BV: Fachwissen(t₁) der Gruppe B, AN: Fachwissen(t₂) der Gruppe A, BN: Fachwissen(t₂) der Gruppe B

F) Ergebnisse der Pilotierung zur Selbstwirksamkeitserwartung

Pretest				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SV1	3.29	1.09	3.43	1.03
SV2	2.66	1.05	2.60	1.13
SV3	2.81	1.12	2.79	1.26
SV4	3.48	.97	3.70	.93
SV5	2.09	1.05	2.14	1.12
Gesamt	2.87	.85	2.93	.88
Posttest				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SN1	3.64	1.12	3.89	1.05
SN2	3.05	1.03	3.11	1.21
SN3	2.89	1.27	3.16	1.30
SN4	3.41	1.25	3.63	1.04
SN5	2.74	1.31	2.57	1.46
Gesamt	3.15	1.05	3.27	

Anmerkungen. *SV* = Selbstwirksamkeitserwartung(t_1), *SN* = Selbstwirksamkeitserwartung(t_2), 1 = stimme nicht zu, 5 = stimme zu; *M* = Mittelwert, *SD* = Standardabweichung.

Ergebnisse der Pilotierung zu den Selbstkonzepten

Physikbezogenes Selbstkonzept				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKP1	3.19	1.00	3.08	1.08
SKP2	3.03	1.23	3.16	1.27
SKP3	3.29	1.48	3.35	1.38
SKP4	2.36	1.11	2.37	.99
SKP5	3.01	1.10	3.06	1.26
SKP6	3.10	.97	3.27	1.05
SKP7	3.53	1.12	4.27	.90
Gesamt	3.08	.79	3.22	.85

Computerbezogenes Selbstkonzept				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKC1	3.23	.90	3.76	1.10
SKC2	3.78	1.06	3.92	1.04
SKC3	2.85	.98	3.30	1.27
SKC4	2.77	1.11	3.52	1.09
SKC5	2.67	1.18	3.30	1.19
SKC6	3.31	1.17	3.76	1.19
SKC7	3.11	1.35	3.51	1.20
Gesamt	3.12	.89	3.58	.91

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu; M = Mittelwert, SD = Standardabweichung.

G) Ergebnisse der Pilotierung zum Flow-Erleben

Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
FL1	3.14	1.32	3.52	1.13
FL3	3.06	1.40	3.73	1.01
FL4	2.83	1.36	3.39	1.30
FL5	3.30	1.33	3.15	1.31
FL6	3.71	1.33	3.18	1.41
FL7	3.18	.97	3.48	1.26
FL8	3.12	1.23	2.87	1.26
FL9	3.41	0.97	2.87	1.20
FL10	3.10	1.16	3.06	1.34
Gesamt	3.21	.59	3.24	.83

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu; *M* = Mittelwert, *SD* = Standardabweichung.

H) Ergebnisse der Erhebung zum Fachwissen

Pretest									
Item	EX		KG		Item	EX		KG	
	M	SD	M	SD		M	SD	M	SD
AV1	.19	.39	.24	.42	BV1	.28	.45	.37	.48
AV2	.44	.50	.30	.45	BV2	.27	.45	.22	.41
AV3	.47	.50	.46	.50	BV3	.34	.48	.39	.48
AV4	.40	.49	.36	.48	BV4	.52	.50	.35	.47
AV5	.34	.48	.28	.45	BV5	.35	.48	.22	.41
AV6	.19	.39	.19	.39	BV6	.05	.22	.06	.24
AV7	.36	.48	.34	.47	BV7	.28	.45	.27	.44
AV8	.34	.48	.29	.45	BV8	.08	.27	.08	.27
AV9	.09	.29	.15	.35	BV9	.06	.23	.07	.25
AV10	.11	.32	.12	.32	BV10	.19	.39	.17	.37
AV11	.43	.50	.35	.48	BV11	.10	.31	.11	.31
AV12	.26	.44	.26	.44	BV12	.34	.48	.30	.45
AV13	.20	.40	.20	.39	BV13	.16	.37	.17	.37
Gesamt	0,31	0,18	0,28	0,17	Gesamt	0,22	0,18	0,19	0,18

Posttest									
Item	EX		KG		Item	EX		KG	
	M	SD	M	SD		M	SD	M	SD
AN1	.62	.49	.49	.50	BN1	.63	.49	.54	.50
AN2	.48	.50	.36	.47	BN2	.39	.49	.34	.47
AN3	.51	.50	.43	.49	BN3	.62	.49	.61	.48
AN4	.52	.50	.45	.49	BN4	.52	.50	.47	.50
AN5	.29	.46	.33	.47	BN5	.52	.50	.62	.48
AN6	.25	.44	.22	.41	BN6	.11	.32	.13	.33
AN7	.35	.48	.34	.47	BN7	.37	.49	.42	.49
AN8	.45	.50	.39	.49	BN8	.13	.34	.16	.36
AN9	.19	.40	.26	.44	BN9	.03	.18	.06	.24
AN10	.17	.38	.13	.33	BN10	.09	.28	.11	.30
AN11	.21	.41	.31	.46	BN11	.13	.34	.13	.33
AN12	.35	.48	.30	.45	BN12	.27	.45	.41	.49
AN13	.36	.48	.23	.42	BN13	.22	.41	.19	.39
Gesamt	.40	.24	.34	.22	Gesamt	.33	.20	.32	.21

Anmerkungen. EX = Experimenttaggruppe. KG = Kontrollgruppe, M = Mittelwert, SD = Standardabweichung.

AV: Fachwissen(t₁) der Gruppe A, BV: Fachwissen(t₁) der Gruppe B, AN: Fachwissen(t₂) der Gruppe A, BN: Fachwissen(t₂) der Gruppe B

I) Ergebnisse der Pilotierung der Selbstwirksamkeitserwartung

Pretest				
	Experimentalgruppe		Kontrollgruppe	
Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SV1	3.37	1.04	3.37	1.07
SV2	2.76	1.12	2.66	1.12
SV3	2.94	1.14	2.99	1.14
SV4	3.32	1.07	3.31	1.09
SV5	3.00	1.21	2.99	1.20
Gesamt	3.08	.86	3.06	.86

Posttest				
	Experimentalgruppe		Kontrollgruppe	
Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SN1	3.57	1.00	3.52	1.01
SN2	2.96	1.00	2.94	1.05
SN3	3.07	1.04	3.09	1.03
SN4	3.36	1.06	3.31	1.05
SN5	3.08	1.14	2.96	1.19
Gesamt	3.25	.80	3.16	.82

Anmerkungen. SV = Selbstwirksamkeitserwartung(t_1), SN = Selbstwirksamkeitserwartung(t_2), 1 = stimme nicht zu, 5 = stimme zu; *M* = Mittelwert, *SD* = Standardabweichung.

J) Ergebnisse der Pilotierung der Selbstkonzept

Physikbezogenes Selbstkonzept				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKP1	3.19	0.98	3.11	1.00
SKP2	3.15	1.21	3.00	1.19
SKP3	3.30	1.36	3.32	1.29
SKP4	2.54	1.16	2.40	1.10
SKP5	3.21	1.26	3.02	1.28
SKP6	3.13	0.99	3.05	1.05
SKP7	2.85	1.32	2.93	1.36
Gesamt	3.05	0.78	2.98	.79

Computerbezogenes Selbstkonzept				
Item	Experimentalgruppe		Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SKC1	3.44	1.02	3.53	1.03
SKC2	3.55	1.30	3.54	1.20
SKC3	3.15	1.02	3.21	1.07
SKC4	3.00	1.21	3.10	1.13
SKC5	2.91	1.11	2.97	1.04
SKC6	3.74	1.24	3.80	1.19
SKC7	3.22	1.30	3.40	1.23
Gesamt	3.29	.93	3.36	.85

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu; *M* = Mittelwert, *SD* = Standardabweichung.

K) Ergebnisse der Erhebung zum Flow-Erleben

Item	Experimentalgruppe		Kontrollgruppe	
	M	SD	M	SD
FL1	3.32	1.20	3.28	1.01
FL3	3.61	1.34	3.52	1.10
FL4	3.30	1.27	3.12	1.18
FL5	3.31	1.29	3.06	1.19
FL6	3.74	1.20	3.29	1.19
FL7	3.37	1.18	3.05	1.22
FL8	3.33	1.11	3.07	1.13
FL9	3.45	1.18	3.27	1.09
FL10	3.10	1.21	2.60	1.09
Gesamt	3.39	0.81	3.14	0.71

Anmerkungen. 1 = stimme nicht zu, 5 = stimme zu; *M* = Mittelwert, *SD* = Standardabweichung.

Eidesstattliche Erklärung

Hiermit erkläre ich durch meine Unterschrift an Eides statt:

1. Die eingereichte Dissertation mit dem Titel *Wirkungen des Flow-Erlebens in einem Serious Game auf die Veränderung von Fachwissen und Selbstwirksamkeitserwartung* habe ich selbstständig und ohne unzulässige fremde Hilfe verfasst. Hierbei habe ich weder Textstellen von Dritten oder aus eigenen Prüfungsarbeiten, noch Grafiken oder sonstige Materialien ohne Kennzeichnung übernommen.
2. Es sind ausschließlich die von mir angegebenen Quellen und Hilfsmittel verwendet worden.
3. Sämtliche wörtliche und nicht wörtliche Zitate aus anderen Werken sind gemäß den wissenschaftlichen Zitierregeln kenntlich gemacht.
4. Die von mir vorgelegte Arbeit ist bisher noch nicht, auch nicht teilweise, veröffentlicht worden.
5. Die von mir vorgelegte Arbeit ist bisher noch in keiner Form als Bestandteil einer Prüfungs-/Qualifikationsleistung vorgelegt worden.
6. Die von mir eingereichte Dissertation habe ich unter Beachtung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis erstellt.
7. Die geltende Promotionsordnung ist mir bekannt.
8. Über die Bedeutung und die strafrechtlichen Folgen einer falschen eidesstattlichen Erklärung gemäß § 156 StGB bin ich mir bewusst.
9. Ich erkläre an Eides statt, dass meine Angaben der Wahrheit entsprechen und ich diese nach bestem Wissen und Gewissen gemacht habe.

Karlsruhe, 03.08.2023

.....