

Planspielbasierte Lehrveranstaltungen – Gelingensbedingungen und Evaluation von Rahmenbedingungen, Reflexion und Einflussfaktoren

Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie,

genehmigt durch die

Fakultät für Humanwissenschaften

der Otto-von-Guericke-Universität Magdeburg

von Tobias Alf

geb. am 05.09.1987 in Oberndorf a.N.

Gutachter: Prof. Dr. Philipp Pohlenz

Gutachterin: Prof. Dr. Silke Geithner

Eingereicht am: 06.08.2024

Verteidigung der Dissertation am: 13.05.2025

Inhalt

1	Einleitung	1
2	Evaluation von Lehre an Hochschulen	4
2.1	Evaluation – eine kurze Einführung.....	4
2.2	Evaluation: Anspruch und empirische Praxis.....	7
2.3	Evaluation an Hochschulen.....	9
2.4	Lehrveranstaltungsevaluation	10
3	Planspiele und ihr Einsatz in der Hochschullehre	13
3.1	Planspiel: Abgrenzung und Definition.....	13
3.2	Lerntheoretische und didaktische Einordnung von Planspielen	17
3.3	Warum Planspiele? Versuch einer Begründung	19
3.4	Typologisierung von Planspielen.....	22
4	Evaluation planspielbasierter Lehrveranstaltungen: Beschreibung des Fragebogens.....	26
4.1	Anknüpfung an unterschiedliche Forschungsstränge	26
4.2	Der ZMS-Fragebogen	28
4.3	Explorative Faktorenstruktur	31
4.4	Reliabilität und Validität.....	34
5	Durchgeführte Studien	41
5.1	Gelingensbedingungen von Planspielveranstaltungen – Ein Systematic Literature Review	45
5.2	Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre.....	68
5.3	The Role of Reflection in Learning with Simulation Games – A Multi-Method Quasi Experimental Research	86
5.4	Impact Factors on Learning and Satisfaction in Simulation Game-based Courses in Higher Education.....	122
6	Diskussion der Ergebnisse und Beitrag zur Forschung	150
6.1	Gelingensbedingungen Planspielbasierter Lehrveranstaltungen	152
6.2	Rahmenbedingungen: Ein Vergleich zwischen Online- und Präsenzlehre	152
6.3	Strukturierte Reflexionsaufgaben	154
6.4	Einflussfaktoren auf Lernen und Zufriedenheit	155
6.5	Fragebogenentwicklung und Ansätze zur Weiterentwicklung	156

7	Exkurs: Kritische Reflexion planspiel- und erfahrungsorientierter Lehre ...	158
8	Fazit und Ausblick	162
8.1	Empfehlungen für die praktische Arbeit am ZMS.....	163
8.2	Ansätze für die weitere Forschung	166
	Literatur (Rahmentext).....	170
9	Anhang.....	179
	Angenommene Faktorenstruktur Pretest.....	179

Abbildungen (Rahmentext)

Abbildung 1: Aufbau der Arbeit.....	3
Abbildung 2: Modell der akademischen Lehre nach Reinmann, 2016 (eigene Darstellung).....	18
Abbildung 3: Anknüpfung an Forschungsstränge	28
Abbildung 4: Evaluationsmodell	30
Abbildung 5: Literaturstudie und daraus abgeleitete empirische Artikel	42
Abbildung 6: Beitrag zur Forschung	151

Tabellen (Rahmentext)

Tabelle 1: Zeitliche Perspektive und Formen von Evaluation.....	7
Tabelle 2: Typologisierung von Planspielen nach Klabbers (2018).....	23
Tabelle 3: Planspiele nach Typisierung	25
Tabelle 4: Explorative Faktorenanalyse (N=3067).....	33
Tabelle 5: Faktorenstruktur, Deskriptive Statistik, Reliabilität (N=3067).....	36
Tabelle 6: Konfirmatorische Faktorenanalyse: Fit-Indikatoren und Schwellenwerte (N=3067).....	38
Tabelle 7: Konvergente Validität, Vergleich mit MEEGA+.....	39
Tabelle 8: Konvergente Validität: Vergleich Lehrende und Studierende	40

1 Einleitung

Diese Arbeit ist eng verknüpft mit dem Zentrum für Managementsimulation (ZMS) an der Dualen Hochschule Baden-Württemberg (DHBW) Stuttgart. Dort werden für alle Studiengänge der Fakultät Wirtschaft und Gesundheit an der DHBW Stuttgart Planspiele und weitere interaktive Lehrmethoden angeboten. Möchte ein Studiengang Lehrinhalte nicht nur theoretisch vermitteln, sondern für Studierende erfahrbar machen, so ist das ZMS ein häufig konsultierter Ansprechpartner. Mit über 30 unterschiedlichen Planspielen und entsprechenden didaktischen Konzepten können sehr viele Themenbereiche dieser Fakultät bedient werden. Neue Themen und Planspiele werden stets bedarfsorientiert aufgegriffen und ins Portfolio aufgenommen.

2008 gegründet, wurde das Angebot des ZMS im Laufe der Jahre vermehrt nachgefragt. Im Durchschnitt besuchen alle Studierenden der Fakultät Wirtschaft und Gesundheit das ZMS im Laufe ihres Bachelorstudiums drei Mal.

Bereits seit seiner Gründung verfolgte das ZMS die Idee von planspielbasierter Lehre einerseits und anwendungsorientierter Forschung zu Lehr-Lernprozessen mit Planspielen andererseits. Die bisherige Forschungs- und Publikationstätigkeit von Mitarbeitenden am ZMS lässt sich grob in drei Stränge gliedern:

- Zunächst finden sich Beiträge mit eher konzeptionellem Charakter, in denen die Praxis und Erfahrung am ZMS beispielhaft beschrieben und reflektiert wird. So geben Zürn, Trautwein und Freese Empfehlungen zur curricularen Integration von Planspielen in Wirtschaftsstudiengänge und knüpfen hier stark an die Konzeption des ZMS an (Zürn et al., 2023). Ein zweites Beispiel hierfür ist Bartschat et al. (2018). Hier wird das Konzept des ZMS für Planspielseminare für Erstsemester vorgestellt und gesammelte Erfahrungen praxisorientiert reflektiert.
- Des Weiteren finden sich vermehrt Veröffentlichungen zum Themenkomplex Facilitation/Debriefing von Planspielen. Diese fokussieren die Rolle und das Selbstverständnis von Lehrenden in planspielbasierten Lehrveranstaltungen. Exemplarisch sei hier die Entwicklung eines Ansatzes (Design of an Impuls-Debriefing-Spiral) genannt, der Debriefing

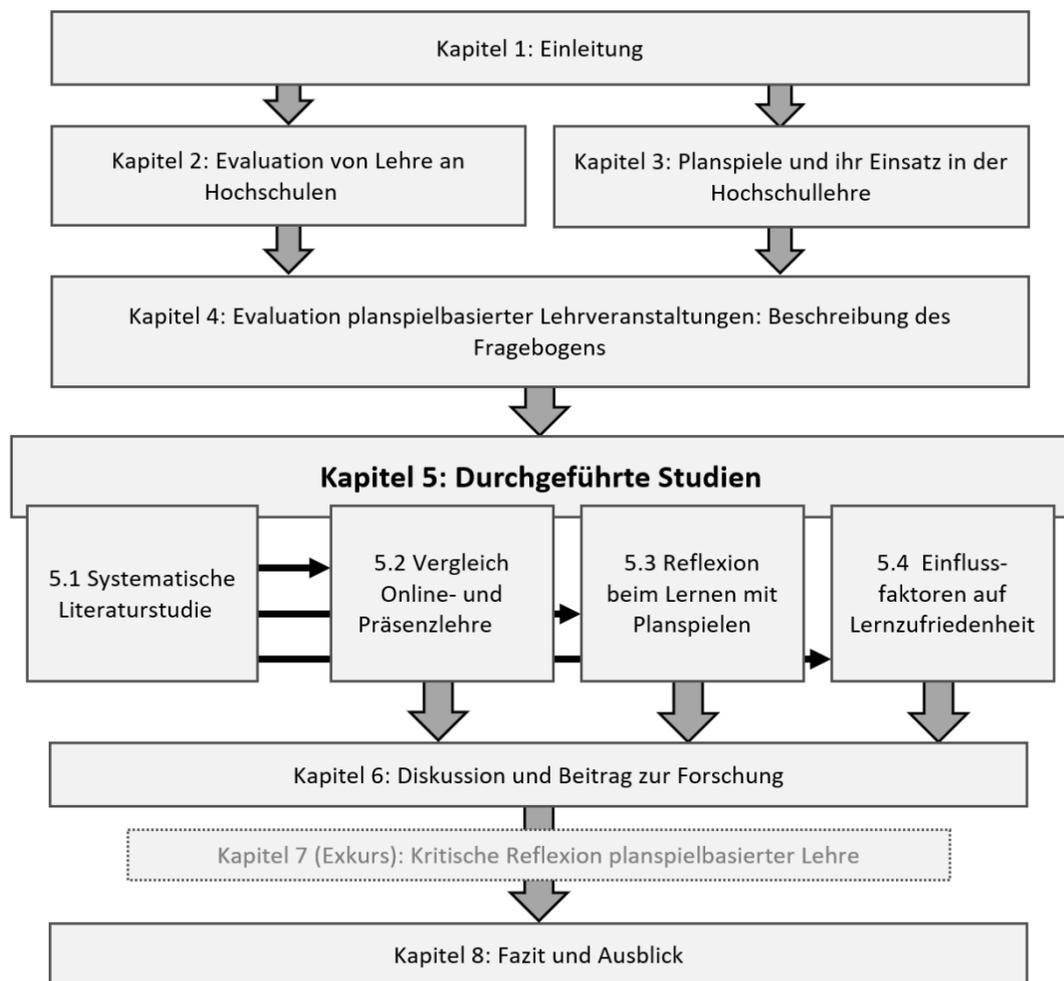
nicht als abgeschlossene Einheit am Ende eines Planspiels versteht, sondern als fortwährende Begleitung der Lernenden während des Planspiels (Schwägele et al., 2021).

- Als dritter Strang ist das Thema Lerntransfer im Kontext von Planspielen zu nennen, das eng mit der Dissertation von Sebastian Schwägele (ehemaliger Leiter und Wissenschaftlicher Mitarbeiter am ZMS) und angrenzenden Veröffentlichungen verbunden ist. Schwägele beleuchtet das Thema mit qualitativen Interviews und betont einen Transfer in zwei Schritten: Vorher Gelerntes muss im Planspiel praktisch angewendet werden und im Planspiel Gelerntes kann später im eigentlichen Handlungsfeld angewendet werden (Schwägele, 2015).

Die bisherigen Veröffentlichungen haben eher konzeptionellen/theoretischen Charakter oder arbeiten mit qualitativen Methoden. Obwohl es sich aufgrund der Vielzahl an Planspielveranstaltungen anbietet, wurde am ZMS bislang wenig mit quantitativen Methoden gearbeitet. Genau dies ist der praktische Ansatzpunkt dieser Arbeit. Die am ZMS durchgeführten Planspielseminare sollen mit vorwiegend quantitativen Methoden evaluiert werden.

Den Kern dieser kumulativen Dissertation bilden vier veröffentlichte bzw. zur Veröffentlichung eingereichte Artikel. Der erste Artikel ist eine systematische Literaturanalyse. Die weiteren drei Artikel greifen in der Literaturanalyse identifizierte Forschungsfragen auf. Diese Artikel sind in Kapitel 5 Durchgeführte Studien enthalten. Der Aufbau der gesamten Arbeit ist in Abbildung 1 dargestellt. Zu Beginn stehen drei Kapitel, die einleitende, übergeordnete Fragen behandeln: Kapitel 2 gibt einen Überblick zum Thema Evaluation von Lehre an Hochschulen. Kapitel 3 behandelt die Rolle von Planspielen in der Hochschullehre. Kapitel 4 stellt den in allen empirischen Studien verwendeten Fragebogen zur Evaluation von Planspielen in der Hochschullehre vor. Nach der Präsentation der veröffentlichten Artikel in Kapitel 5 folgt eine übergreifende Diskussion der Ergebnisse und deren Beitrag zur Forschung (Kapitel 6). Bevor die Arbeit mit einem Fazit (Kapitel 8) schließt, wird in Kapitel 7 als Exkurs eine kritische Auseinandersetzung mit Planspielen und erfahrungsorientierten Lehrmethoden allgemein, eingeschoben.

Abbildung 1: Aufbau der Arbeit



Eine kumulative Dissertation bringt die Schwierigkeit mit sich, bereits editierte und veröffentlichte Artikel in eine einheitliche Form zu bringen, möglichst ohne ihren jeweiligen Charakter zu beeinflussen. Daher bedarf es einiger Vorbemerkungen zur Lesbarkeit dieser Arbeit. Wie üblich findet sich am Ende ein Literaturverzeichnis. Dieses bezieht sich auf den Rahmentext, also die einleitenden Kapitel 1 bis 4 sowie die abschließenden Kapitel 6 bis 8. Die in Kapitel 5 dargestellten Artikel enthalten ihr jeweiliges Literaturverzeichnis, wie veröffentlicht, jeweils an ihrem Ende. Das gleiche gilt für Abbildungs- und Tabellenverzeichnisse. Das eingangs gedruckte Tabellen- und Abbildungsverzeichnis bezieht sich auf den Rahmentext. Die Abbildungen und Tabellen in Kapitel 5 beginnen jeweils bei eins und entsprechen somit ihrer ursprünglichen Nummerierung, wie veröffentlicht. Ebenso wird mit Anhängen verfahren. Die Anhänge der publizierten Artikel finden sich direkt an deren Ende,

während sich der Anhang ganz am Ende des Gesamttextes auf dem Rahmentext bezieht. Pointiert könnte man sagen: Die in Kapitel 5 enthaltenen Artikel sollten wie Aufsätze in Fachzeitschriften gelesen werden, der Rahmentext eher wie eine Monografie.

2 Evaluation von Lehre an Hochschulen

Bevor sich diese Arbeit dem Thema Planspiele in der Hochschullehre und deren Evaluation widmet, steht eine kurze Einführung zum Thema Evaluation an Hochschulen. Zunächst wird der Begriff Evaluation eingeführt. Kapitel 2.2 lotet die Diskrepanz zwischen (theoretischem) Anspruch an Evaluation und deren Durchführbarkeit in der Praxis aus. Das nächste Unterkapitel widmet sich verschiedenen, an Hochschulen typischen Bewertungs- und Evaluationsverfahren, während Kapitel 2.4 das für diese Arbeit relevante Feld der Lehrveranstaltungsevaluation behandelt.

2.1 Evaluation – eine kurze Einführung

Das Wort Evaluation hat unter anderem Wurzeln in der lateinischen (valēre 'stark sein, vermögen, geeignet, imstande, wert sein'), französischen (valoir 'wert sein')¹ und auch englischen (value 'Wert') Sprache (Gnahn & Quilling, 2019, S. 99). Die Bedeutungen überschneiden sich sehr stark: Einen Wert haben, von Wert sein. Evaluation meint heute den Wert einer Sache zu beurteilen, kurz: zu bewerten.

Die Deutsche Gesellschaft für Evaluation e.V. definiert Evaluation in ihrem Glossar wie folgt (Deutsche Gesellschaft für Evaluation e.V., o.D.):

Die systematische Untersuchung von Nutzen und/oder Güte eines Gegenstands (Evaluationsgegenstand) auf Basis von empirisch gewonnenen Daten. Impliziert eine Bewertung anhand offengelegter Kriterien für einen bestimmten Zweck.

¹ „Evaluation“, in: DWDS – Digitales Wörterbuch der deutschen Sprache, hrsg. v. d. Berlin-Brandenburgischen Akademie der Wissenschaften, <<https://www.dwds.de/wb/Haus>>, abgerufen am 27.09.2023.

Auch die Definition der Deutschen Gesellschaft für Evaluation hebt den Charakter der Bewertung hervor, spezifiziert jedoch, *wie* diese Bewertung vorgenommen wird: Durch systematische Untersuchung, die sich auf einen definierten Gegenstand (Evaluationsgegenstand) konzentriert. Hierfür werden empirische Daten verwendet und sowohl die Bewertungskriterien als auch der Zweck der Bewertung sind transparent.

Evaluation ist eng verknüpft mit empirischer Sozialwissenschaft (Kromrey, 2005) und geht als solche weit über die „Selbstreflexion der alltäglichen Praxis“ (Böttcher et al., 2006, S. 9) hinaus. Evaluation arbeitet mit den theoretischen und methodischen Standards der empirischen Forschung und nicht mit Alltagsbewertungen. Wenn man einen Unterschied zwischen Forschung und Evaluation beschreiben möchte, so besteht dieser nicht in der methodischen Herangehensweise, sondern in den Zielen der empirischen Untersuchung. Während in der Forschung das zweckfreie Erkenntnisinteresse und die Theorieentwicklung im Vordergrund stehen, möchte Evaluation ein konkretes Programm oder eine konkrete Organisation bewerten. Im Vordergrund steht eine praktische Frage, die beantwortet werden soll (Kromrey, 2005; Pohlenz, 2008a, S. 27). Ganz trennscharf lassen sich Evaluations- und Forschungsparadigma jedoch nicht immer unterscheiden.

Stockmann (2006, S. 30) beschreibt vier eng miteinander verknüpfte Funktionen von Evaluation: Sie soll **Erkenntnisse** liefern. Der Evaluationsgegenstand soll anhand transparenter Kriterien besser verstanden werden, damit Entscheidungen „auf eine rationale Grundlage gestellt werden“ (ebd.) können. Nicht immer stehen Erkenntnisse zur Entscheidungsfindung im Vordergrund, häufig geht es im Kontext von Evaluation um **Kontrolle**: Es soll festgestellt werden, „ob die in der Planung festgelegten Ziele erreicht wurden“ (ebd.). Auch die (Weiter-)**Entwicklung** von Programmen oder Lehrveranstaltungen können Zweck der Evaluation sein. Anhand gewonnener Erkenntnisse kann im Dialog mit Stakeholdern über die Art der künftigen Fortführung/Veränderung/Verbesserung des Evaluationsgegenstandes gesprochen werden. Eine weitere Funktion von Evaluation ist die **Legitimation**. Anhand der empirischen Datenbasis können Programme/Maßnahmen gegenüber Stakeholdern (häufig Finanziers) begründet werden.

Bei der Funktion Erkenntnis sieht Kromrey (2005) eine starke Überschneidung zwischen Forschungsparadigma (Theorieentwicklung) und Evaluationsparadigma

(Bewertung mit praktischem Nutzen): Neben dem eigentlichen Evaluationszweck könnten auch grundlagenwissenschaftliche Ziele verfolgt werden. Die Evaluation der Lehre mit Planspielen am ZMS an der DHBW Stuttgart verfolgt vorwiegend den Zweck Erkenntnis und versucht so, einen Mittelweg zwischen Evaluationszweck und wissenschaftlichen Ziele zu beschreiten. Es soll verständlich werden, welche Faktoren für das Lernen mit Planspielen ausschlaggebend sind. Sodann dienen die Erkenntnisse der Weiterentwicklung von didaktischen Konzepten am ZMS und ggf. auch der Legitimation planspielbasierter Lehre gegenüber der Hochschulleitung. Der Aspekt Kontrolle wird mit der durchgeführten Evaluation weniger verfolgt. So sind Auswertungen auf Ebene einzelner Lehrender explizit nicht Teil des Evaluationsprojektes.

Evaluation wird in der Literatur häufig zusammen mit Qualitätsmanagement gedacht und diskutiert (Böttcher et al., 2006, S. 7; Hense et al., 2019, S. 7; Pohlenz, 2008b; Stockmann, 2006). Obwohl die beiden Konzepte über Schnittmengen verfügen, zeigen sich doch auch wesentliche Unterschiede: Qualitätsmanagement (QM) ist ursprünglich ein Managementkonzept betriebswirtschaftlich agierender Unternehmen und hat eine „klare Produzenten-Kunden-Beziehung“ im Blick. QM arbeitet daher nicht „nur“ im Bereich der Informationsbeschaffung und Auswertung, sondern schließt auch die Umsetzung daraus folgender Unternehmensentscheidungen mit ein (Stockmann, 2006, S. 35). Evaluation hingegen meint eine systematische und methodisch abgesicherte Bewertung eines definierten Gegenstandes (Böttcher et al., 2006, S. 9). Somit arbeitet Evaluation mit Daten und zieht aus diesen Schlussfolgerungen (Bewertungen) – ohne jedoch in daraus womöglich resultierendes Managementhandeln eingebunden zu sein.

Für die praktische Durchführung von Evaluationen werden in der Literatur, je nach Zielsetzung und zeitlicher Einbeziehung der Evaluation, verschiedene grundsätzliche Formen beschrieben und diskutiert. Häufig wird zwischen formativer und summativer Evaluation unterschieden (Kromrey, 2005; Stockmann, 2006). Formative Evaluation meint gestaltende, prozessbegleitende Evaluation. Im Verlauf der Evaluation werden Ergebnisse direkt rückgekoppelt und das Programm ggf. entsprechend angepasst. Formative Evaluation muss daher in der Planungsphase oder spätestens in der Durchführungsphase des zu evaluierenden Projektes ansetzen, da ansonsten die Evaluationsergebnisse nicht mehr in die Gestaltung des Projekts einfließen könnten. Summative Evaluation hingegen

meint bilanzierende, ergebnisorientierte Evaluation. Die gesammelten Ergebnisse werden nicht direkt für eine mögliche Anpassung eines Programms verwertet, sondern zur rückblickenden Beurteilung. Aus dieser können dann selbstverständlich Schlüsse für künftige Gestaltungen ähnlicher Programme gezogen werden. Summative Evaluation setzt daher üblicherweise frühestens in der Durchführungsphase an.

Tabelle 1: Zeitliche Perspektive und Formen von Evaluation²

Programm-/Projektphase	Zeitliche Perspektive	Mögliche Form der Evaluation
Planung	Vorausblickend (ex-ante)	Formativ, begleitend
Durchführung	On-going	Formativ oder summativ
Abschluss/Rückblick	Rückblickend (ex-post)	Summativ, bilanzierend

2.2 Evaluation: Anspruch und empirische Praxis

Die in der Literatur gestellten Ansprüche an Evaluation sind sehr hoch. Evaluation soll kausale Zusammenhänge darstellen, etwa, dass die Einstellungsänderung der Teilnehmenden an Programm x tatsächlich auf Programm x zurückzuführen ist, idealerweise anhand von randomized controlled trials (RCT) (Müller, 2017; Spinath & Seifried, 2018). Hierfür wäre es notwendig, dass die Teilnehmenden von Programm x zufällig ausgewählt werden und sie eine hinreichend große Stichprobe umfassen. Mit einer ebenfalls zufällig ausgewählten Kontrollgruppe müsste die Wirkung von Programm x überprüft werden. Zudem wäre durch ein Pre/Post-Design auszuschließen, dass die womöglich gefundenen Effekte durch Reifung oder durch Lerneffekte aus dem Pre-Test eingetreten sind. Für die Evaluation von Hochschullehre sind diese Anforderungen in der Praxis kaum umzusetzen (Pohlenz, 2008b, S. 70). Beispielsweise scheitern sie häufig bereits an der Randomisierung. Studierende bewerben sich aus Interesse für ein Fach, nehmen aus Interesse an bestimmten Lehrveranstaltungen teil oder sie müssen

² Die Gestaltung der Tabelle erfolgte in Anlehnung an Stockmann, 2006.

diese, um ein Zertifikat zu erhalten, schlicht besuchen; was eine zufällige Auswahl praktisch unmöglich macht. Störfaktoren können somit nicht sicher ausgeschlossen werden. Zudem sind Lehr-Lern-Prozesse sehr vielschichtig. Lehner und Sohm (2021, S. 349) warnen davor, verkürzt „nach der ‚richtigen‘ didaktischen Methode“ zu suchen und plädieren dafür, mit einer qualitativen Perspektive auf den Lehr-Lern-Prozess zu blicken. Als ein Problem der Lehrveranstaltungsevaluation benennen Spinath und Seifried (2018, 157.f) das Kriterienproblem. Es sei schlicht unklar, was genau gute Hochschullehre sei und welche Kriterien evaluiert werden sollten. In der schulischen Unterrichtsforschung wurde das Konzept von Oberflächen- und Tiefenstrukturen des Unterrichts entwickelt (Decristan et al., 2020; Lehner & Sohm, 2021, S. 346). Oberflächenstrukturen meint jene Merkmale von Unterricht, die leicht beobachtbar sind (Anordnung des Klassenraums, Methoden, Unterrichtsmaterialien etc.). Tiefenstrukturen hingegen meint die qualitative Auseinandersetzung mit dem Lerngegenstand, die nicht ohne Weiteres beobachtbar ist. Empirisch gemessen und verglichen werden, da leichter beobachtbar/operationalisierbar, zumeist Oberflächenstrukturen. Es ist jedoch davon auszugehen, dass wesentlich die Tiefenstrukturen des Unterrichts (z.B. Konzentration, Kommunikation) zum Lehr-Lernprozess beitragen. Für den Bereich der Hochschullehre dürfte dies nicht anders sein. Selbst anspruchsvolle experimentelle Vergleichsstudien könnten demnach übersehen, welche Elemente im Lehr-Lernprozess ausschlaggebend sind. Genau dieser Punkt wird an Studien, die kausale Zusammenhänge untersuchen, kritisiert. Wird ein Effekt zwischen den Gruppen gefunden, könnte dennoch unklar bleiben (black box), wie dieser genau entstanden ist (Müller, 2017, S. 215). Neben dem methodischen Goldstandard (RCT) werden daher Formen der Evaluation vorgeschlagen, die geeignet sind theoretisch angenommene Prozesse sichtbar zu machen: theoriegeleitete Evaluation (Müller, 2017, S. 215; Pohlenz, 2008a, S. 42, 2008b). Hierbei wird eine Theorie des Programms formuliert und operationalisiert. Anhand hypothesentestender Verfahren können die theoretisch angenommenen Zusammenhänge und Wirkungsweisen des Programms getestet und beurteilt werden. Auch in dem hier dargestellten Forschungsvorhaben ist die Umsetzung einer RCT-Studie aus den oben dargestellten Gründen praktisch nicht möglich. Beispielsweise gehören die Planspiele zum Studienprogramm der Studierenden, was es praktisch unmöglich macht, eine passende Kontrollgruppe zu finden. Zudem könnte kaum ausgeschlossen werden, dass Studierende die Inhalte in anderen Veranstaltungen gelernt haben, die sie direkt vorher oder

nachher besuchen. Der im Rahmen dieser Arbeit entwickelte Fragebogen (siehe Kapitel 4) ist ein typisches Beispiel für theoriegeleitete Evaluation. Anhand operationalisierter Variablen wird versucht, das „Programm“ Lehrveranstaltungen mit Planspielen abzubilden. Anhand hypothesenprüfender Verfahren kann anschließend getestet werden, inwiefern die angenommenen Zusammenhänge in den Daten sichtbar werden. In den empirischen Studien (Kapitel 5) werden die Evaluationsdaten jedoch ganz unterschiedlich in das jeweilige Forschungsdesign eingebunden und teilweise mit anderen Daten oder Erhebungsinstrumenten verbunden.

2.3 Evaluation an Hochschulen

Evaluation meint die strukturierte und methodisch abgesicherte Bewertung eines bestimmten Gegenstandes. Der zu definierende Gegenstand ist keineswegs nur gegenständlich zu denken. Es können Konzepte, Strukturen, Prozesse oder Ergebnisse evaluiert werden (Gnahn & Quilling, 2019). Für Evaluationsvorhaben im Bildungsbereich lassen sich nach Gnahn und Quilling (2019, S. 101) folgende Evaluationsgegenstände unterscheiden:

- Bildungssystem als Ganzes (Systemevaluation)
- Evaluation von Bildungsgesetzen (Gesetzevaluation)
- Evaluation von Bildungseinrichtungen (Einrichtungsevaluation)
- Evaluation von Bildungsveranstaltungen (Veranstaltungsevaluation)
- Evaluation einzelner Aspekte von Veranstaltungen oder Einrichtungen

Spezifisch für Hochschulen beschreiben Schmidt und Pohlenz (2019, 137ff.) drei Evaluationsgegenstände bzw. Handlungsfelder, in denen sich unterschiedliche Praxen der Evaluation etabliert haben: Forschung, Lehre und den Wissenschaft unterstützenden Bereich.

- In der **Forschung** haben sich Peer-Bewertungssysteme durchgesetzt, zumeist ohne, dass explizit von Evaluation gesprochen wird. Projekt- und Drittmittelgelder werden in der Regel nach „umfangreichen Bewertungsverfahren durch Fachkolleginnen und Fachkollegen“ (ebd.) vergeben. Ähnliches gilt für Veröffentlichungen in Fachzeitschriften, bei denen die Bewertung der wissenschaftlichen Arbeit durch Peer-Review-Verfahren erfolgt.
- Für den Bereich **Lehre** haben sich seit Beginn der 1990-Jahre Verfahren etabliert, die explizit als Evaluation benannt werden.

Lehrveranstaltungsevaluation untersucht Fragen der Didaktik und der Struktur von Lehrveranstaltungen, während weitere Befragungen retrospektiv das ganze Studium in den Blick nehmen oder nach Gründen für den Studienabbruch suchen. Ein Paradigmenwechsel ergab sich durch die Einführung der Systemakkreditierung, mit der nicht länger Studiengänge evaluiert werden, sondern das Qualitätsmanagement der Hochschulen selbst. Die Bewertung des Systems zielt darauf ab festzustellen, ob das Qualitätsmanagement einer Hochschule so designt ist, dass es die Qualität von Studium und Lehre selbst erfassen und verbessern kann.

- Im **Wissenschaft unterstützenden Bereich** haben sich nach Schmidt und Pohlenz insbesondere der Industrie entlehnte Bewertungsverfahren (Anlehnungen an ISO und Total Quality Management) etabliert. Diese zielen vorwiegend auf die effizientere Gestaltung von Prozessen und eine bessere Kundenorientierung.

Da sich diese Arbeit mit Planspielen in der Hochschullehre auseinandersetzt, konzentriert sich die theoretische Auseinandersetzung im Folgenden auf die Besonderheiten der Evaluation von Lehre.

2.4 Lehrveranstaltungsevaluation

Um zu präzisieren, was für diese Arbeit der relevante Evaluationsgegenstand ist, orientiere ich mich an einer Differenzierung nach Rindermann (2003, S. 304), der zwischen Lehrevaluation und Lehrveranstaltungsevaluation unterscheidet: Lehrevaluation bezieht sich auf das Studium als Ganzes und umfasst sämtliche Lehrveranstaltungen und auch Rahmenbedingungen, wie z.B. die Bibliotheksausstattung. Lehrveranstaltungsevaluation hingegen bezieht sich ausschließlich auf konkrete Lehrveranstaltungen und ist somit ein Teilgebiet der Lehrevaluation. Diese Unterscheidung klingt auch bei Schmidt und Pohlenz (2019) an, wenn sie Evaluation im Bereich Lehre beschreiben: Sie grenzen die Evaluation spezifischer Veranstaltungen von der Evaluation ganzer Studienprogramme ab.

Für die Evaluation von Lehrveranstaltungen sind seit den 1990er Jahren zahlreiche empirische Instrumente (zumeist Fragebögen) entstanden, die sich in ihrer wissenschaftlichen Qualität jedoch stark unterscheiden (Knödler, 2019, S. 114). Häufig beziehen sich deutschsprachige Lehrveranstaltungsinstrumente in ihrer Entwicklung auf das englischsprachige Student Evaluation of Educational

Quality (SEEQ) (MARSH, 1982). Es ist multidimensional aufgebaut (Lernen, Enthusiasmus, Organisation, Interaktion, persönliches Wohlbefinden, Breite, Prüfungen, Aufgaben, Allgemeinbewertung) und empirisch vielfach erprobt. Im deutschsprachigen Raum gilt das Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE) (Rindermann, 2009) als etabliertes und ebenfalls empirisch validiertes Instrument zur Erfassung der Lehrveranstaltungsqualität (Knödler, 2019, S. 114). Es umfasst in mehreren Dimensionen Rahmenbedingungen der Lehre, das Dozierendenverhalten sowie Studierende als Einflussfaktoren und untersucht deren Zusammenwirken auf den Lehrerfolg (Rindermann, 2003). Das Instrument wird umfassend beschrieben und anhand vieler Methoden auf Reliabilität und Validität geprüft (Rindermann, 2009). Das vom Arbeitskreis „Lehrevaluation“ im Fach Psychologie entwickelte Trierer Inventar zur Lehrevaluation (TRIL) (Arbeitskreis "Lehrevaluation" im Fach Psychologie et al., 2002) ist ebenfalls mehrdimensional aufgebaut (Struktur und Didaktik, Anregung und Motivation, Interaktion und Kommunikation, Persönlicher Gewinn durch Veranstaltung, Anwendungsbezug, Referate, Weiteres und Gesamtbeurteilung). Hinsichtlich Reliabilität und Validität werden konkrete Werte genannt und auf Untersuchungen verwiesen (ebd.). Ein etwas jüngeres und umfassend getestetes Instrument legt Knödler (2019) vor, die sich bei der Fragebogenentwicklung insbesondere auf Items stützt, die „direkt auf das Verhalten der Lehrenden zurückgeführt werden können“, um lernunabhängige Effekte zu vermeiden (S. 135). Neben Einflussfaktoren auf das Lernen (z.B. Atmosphäre, Kommunikation, Präsentation der Inhalte etc.) fragt Lehrveranstaltungsevaluation stets auch nach dem Lernerfolg. Genau hier setzt auch eine zentrale Kritik der Lehrveranstaltungsevaluation an. Studierende könnten den ihren Lernerfolg selbst nicht hinreichend bestimmen (Spinath & Seifried, 2018, S. 165). Metaanalysen zu Studien, in denen Lernerfolge über Selbsteinschätzung einerseits und objektivere Daten andererseits erhoben werden, kommen zu keinen einheitlichen Ergebnissen (Clayson, 2009; Spooen et al., 2013). Sichtbar wird ein breites Spektrum, das von sehr starken Zusammenhängen zwischen Selbsteinschätzung und objektiven Daten bis hin zu negativen Zusammenhängen reicht. Spooen et al. (2013) finden tendenziell positive Korrelationen zwischen Selbsteinschätzung und objektiveren Daten, mahnen jedoch eine sehr vorsichtige Interpretation der Ergebnisse an, da die objektiveren Daten in den Studien von sehr unterschiedlicher Qualität seien. Auch Clayson et al. (2009) finden tendenziell positive Korrelationen zwischen

Selbsteinschätzung und objektiveren Daten (allerdings auch negative). Zudem zeigen sie auf, dass die Zusammenhänge kleiner werden, je objektiver die Messdaten erhoben werden und je mehr die Studien Störvariablen statistisch kontrollieren. Auch sie mahnen eine sehr vorsichtige Interpretation der Ergebnisse an: Entscheidend sei nicht ob, sondern unter welchen Umständen ein Zusammenhang gefunden werde. Damit weisen sie darauf hin, dass die Ergebnisse stark von der Art der Studie und der Objektivität der Lernerhebung abhängen. Zudem spielen auch die im jeweiligen Studienfach typische „Korrektheit“ der objektiven Bewertung eine Rolle: z.B. einerseits Mathematik vs. künstlerische Fächer andererseits (Clayson, 2009, S. 26).

Einen Mittelweg zwischen Selbsteinschätzung und objektiven Daten in der Lehrveranstaltungsevaluation gehen das Bielefelder Modell der Lernzielorientierten Evaluation (Bitterer et al.) und die von Raupach entwickelte Lernerfolgsevaluation (EvaSys, o.J.) (als Plug-in in der Evaluationssoftware Evasys enthalten). Obgleich sich die beiden Ansätze in ihrer Konzeption wesentlich unterscheiden, haben sie doch gemeinsam, dass sie bei der Formulierung der Items für die Messung des Lernerfolgs mit sehr konkreten Lernzielen arbeiten. Anhand dieser sollen Studierende ihr eigenes Lernen beurteilen (bzw. im Falle des Bielefelder Modells auch selbst Lernziele formulieren). Anstelle sehr allgemeiner Items wie „Ich verfüge über ein grundlegendes Verständnis als vor der Veranstaltung.“ (aus HILVE, Rindermann, 2009) werden Items formuliert, die sehr konkrete Lernziele abfragen, wie z.B. „Ich kann Literatur mit Hilfe von Software x im Zitationsstil y zitieren“. Je konkreter das Item das eigentliche Lernziel beschreibt, desto verlässlicher können Studierende ihren tatsächlichen Lernerfolg einschätzen, so die Idee der beiden Konzepte. Für die Validierung des Fragebogens wurde diese Idee aufgegriffen. In einer kleineren Stichprobe wurden lernzielorientierte Items abgefragt und deren Ergebnisse anschließend mit dem eher allgemein abgefragten Lernen korreliert (siehe hierzu Kapitel 4.4).

Mit diesem Überblick zur Lehrveranstaltungsevaluation schließt das einführende Kapitel zum Thema Evaluation nun ab. Bevor jedoch das im empirischen Teil dieser Arbeit verwendete Evaluationsinstrument beschrieben und auf Reliabilität und Validität geprüft wird, folgt eine Einführung zu Planspielen in der Hochschullehre, dem zentralen Evaluationsgegenstand dieser Arbeit.

3 Planspiele und ihr Einsatz in der Hochschullehre

Planspiele blicken auf eine sehr lange Tradition zurück. Vorläufer dessen, was wir heute ein Planspiel nennen, sind Kriegsspiele. Die ältesten dieser Kriegsspiele finden sich in China und lassen sich auf ca. 3000 v. Chr. datieren (Schwägele, 2013, S. 18). In Deutschland lassen sich Planspiele seit dem 18. Jahrhundert in der Offiziersausbildung nachweisen (Geuting, 1992, S. 319). Eine Expansion von Planspielen und angrenzenden simulativen Methoden lässt sich seit den 1960er Jahren insbesondere in Nordamerika und Europa beschreiben. In dieser Zeit entsteht auch das „Management Game“. Die Methode wird somit aus dem militärischen Bereich übertragen auf wirtschaftliche und politische Themen (ebd., S. 317). Hinter dem heutigen Begriff Planspiel verbirgt sich nicht nur eine sehr lange historische Entwicklung, sondern auch eine weite thematische und methodische Breite. Im Folgenden wird daher zunächst definiert, was im Rahmen dieser Arbeit unter einem Planspiel verstanden wird. Anhand des Modells der akademischen Lehre nach Reinmann erfolgt im Anschluss eine didaktische Einordnung der Methode in den Kontext Hochschulbildung. Es folgt eine Begründung, warum Planspiele als Lehr-Lernmethode zum Einsatz kommen. Das Kapitel endet mit einer Typologisierung von Planspielen, in der auch die im empirischen Teil dieser Arbeit evaluierten Planspiele eingeordnet werden.

3.1 Planspiel: Abgrenzung und Definition

Um zu beschreiben was ein Planspiel ist, kann es hilfreich sein, zunächst verwandte Begriffe und Konzepte zu beleuchten, um diese von Planspielen abzugrenzen.

Im englischen Term Simulation Game klingt an, dass Planspiele mit **Simulationen** zu tun haben oder gar Simulationen enthalten. Simulationen sind (zumeist) computerbasierte Abbildungen/Modelle der Realität (Campos et al., 2020, S. 3). Duke spezifiziert, dass in Simulationen insbesondere dynamische Prozesse oder dynamisches Verhalten abgebildet werden (Duke & Geurts, 2004, S. 271). Von einem Modell wird beispielsweise auch dann gesprochen, wenn ein Gebäude statisch in Miniaturform nachgebaut wird. Dies wäre jedoch noch keine Simulation. Eine Simulation modelliert Dynamiken, Prozesse, Zusammenhänge oder Interaktionen zwischen mehreren Elementen (Geuting, 1992, S. 33). Planspiele und Simulationen haben gemeinsam, dass sie Dynamiken der realen Welt

nachbilden wollen. Simulationen konzentrieren sich hierbei auf die technische Exaktheit der Nachbildung und haben vorwiegend ein Erkenntnisinteresse. In Planspielen werden Simulationen in das Spieldesign integriert, um Systeme (z.B. Management, Politik) für Lernende erfahrbar zu machen. Sie haben demnach vorwiegend ein didaktisches Interesse, weshalb auch mit sehr abstrakten Simulationen gearbeitet wird (Schwägele, 2013). Somit ist nicht jede Simulation ein Planspiel, aber jedes Planspiel enthält eine Simulation, mit dem Ziel bestimmte Dynamiken für Lernende erfahrbar und reflektierbar zu machen.

Im Kontext von spielerischen Lernmethoden ist häufig von **Serious Games**, Spielen mit ernsthaftem Charakter, die Rede. Dörner et al. (2016, S. 3) bezeichnen Spiele dann als Serious Games, wenn sie neben der Unterhaltung mindestens ein weiteres „ernsthaftes“ Ziel verfolgen (Lernen, Gesundheit, Fitness etc.). Charakteristisch für Serious Games ist, dass nach wie vor ein starker Fokus auf der Unterhaltung liegt und nicht unbedingt Lernen als weiteres Ziel im Mittelpunkt stehen muss. Planspiele sind keineswegs Unterhaltungsspiele. Sowohl in der Entwicklung als auch in der Anwendung werden stets bestimmte (Lern-)ziele verfolgt. Daher kann ein Planspiel auch als Serious Game bezeichnet werden, aber keinesfalls ist jedes Serious Game ein Planspiel.

Gamification ist ein weiterer Begriff aus dem Themenkomplex Spiel und Lernen. Er meint den Gebrauch von Elementen, die charakteristisch für Spiele sind, in einem Nicht-Spiel-Kontext (Deterding et al., S. 11). Sehr häufig findet Gamification im Marketing Anwendung (vgl. Dörner et al., 2016, S. 4). Ein oft anzutreffendes Beispiel ist das Sammeln von Statuspunkten für den wiederholten Konsum von Dienstleistungen. Für regelmäßige Fahrten mit der Deutschen Bahn können beispielsweise Punkte gesammelt und dadurch verschiedene Level erreicht werden³. Typische Spielmechaniken (Punkte sammeln, Level erreichen) werden so in Nicht-Spiel-Situationen integriert, jedoch ohne, dass ein vollwertiges Spiel entsteht. Planspiele sind komplette Spiele und finden explizit in einem Spiel-Kontext statt. Die Überschneidung von Gamification und Planspielen ist somit als eher gering einzuschätzen.

Als **Lernspiele** bezeichnen Plass et al. (2013, S. 709) Spiele, die so designt sind, dass Spielende während ihrer spielerischen Tätigkeit lernen: „game meachnics

³<https://www.bahn.de/service/bahnbonus/status-level>.

must go beyond making a game fun and engaging – they must engage players in meaningful learning activities. The game mechanic becomes an integral part of the learning activity.” Planspiele können in dem Sinne Lernspiele sein, dass Spielende während ihrer Aktivität im Spiel lernen. Während der Spielaktivität etwas zu lernen ist jedoch noch kein hinreichendes Kriterium für ein Planspiel. Genau wie Serious Games können Planspiele Lernspiele sein, jedoch nicht jedes Lernspiel ist auch ein Planspiel.

In Anlehnung an Frederik Barth beschreibt Klabbers **Spiele** („Games“) als temporäre soziale Organisationen, die aus drei miteinander verbundenen Bausteinen bestehen: Actors, Rules und Resources (2018, S. 219). Durch die Übernahme von Rollen, die anhand von Regeln ausgeübt werden, entsteht im Spiel eine „echte“ temporäre soziale Organisation. In dieser temporären sozialen Organisation stehen Ressourcen zur Verfügung, mit denen diese repräsentiert wird. Dies können sprachliche Symbole oder Gesten sein (für das Spiel bedeutsame Wörter oder Bewegungen), aber auch haptische Symbole (z.B. Münzen, Karten, Spielbrett). Durch Rollenübernahme, Regeln und Ressourcen entsteht im Spiel eine „echte“ Kultur, die sich zirkular selbst erhält: Regelgeleitet agieren Spielende im Rahmen ihrer Rollen und nutzen zur Verfügung stehende Ressourcen. Wenn alle Spielenden ihre Rollen kennen und entsprechend der Regeln auf Ressourcen zugreifen können, entwickelt sich eine Sinnhaftigkeit innerhalb des Spiels, entsprechend der die Teilnehmenden agieren und das Spiel zirkulär erhalten („circular organisation“). Die Beschreibungen von Klabbers verweisen auf die Ernsthaftigkeit des Spiels. Lassen sich Menschen auf die temporäre soziale Organisation ein, entwickelt sich eine immanente Sinnhaftigkeit, die nur innerhalb des Spiels zählt: Spielende wetteifern um Punkte, weisen sich gegenseitig auf die Einhaltung von Regeln hin oder sind sich beispielsweise darüber einig, dass grüne Steinchen wertvoller sind als rote. Auch Spiele sind per se noch keine Planspiele, bilden jedoch die Grundlage eines jeden Planspiels: Kriz greift die allgemeine Beschreibung von „Games“ auf und entwickelt eine Definition von Planspielen (2009, S. 560):

„Planspiele beinhalten Akteure, Regeln und Ressourcen (Klabbers 1999). Planspiele sind der Realität angenäherte Modelle, in denen aber immer Menschen als „Mitspieler“ Rollen übernehmen und konkrete Entscheidungen treffen müssen, deren wirklichkeitsrelevante Aus- und Folgewirkungen dann wiederum geprüft werden.“

Damit benennt Kriz zwei weitere wesentliche Aspekte von Planspielen: 1. „Planspiele sind der Realität angenäherte Modelle“ und 2. „deren wirklichkeitsrelevante Aus- und Folgewirkungen“ werden geprüft. Ein entscheidendes Merkmal von Planspielen ist, dass sie Simulationen oder Modelle der Realität enthalten, während dies für Games nicht notwendigerweise der Fall ist. Games können gerade dadurch attraktiv sein, dass sie über keinen oder nur einen sehr geringen Realitätsbezug verfügen. In Planspielen simulieren und erleben Spielende jedoch modellhafte Nachbildungen der Realität. Das Zusammenwirken aus Akteuren, Regel und Ressourcen (Game) ist in Planspielen so angelegt, dass Realität simuliert wird. Die Nachbildung dieser Systeme kann unterschiedlich konkret oder abstrakt sein (W. Kriz & Manahl, 2016, S. 74), bzw. sich auf die komplexitätsreduzierte Abbildung eines Gesamtsystems beziehen oder auf die möglichst realitätsgetreue Abbildung eines Teilsystems (Schwägele, 2015, S. 26). Zudem werden in Planspielen wirklichkeitsrelevante Wirkungen diskutiert, was auf die Reflexion (Debriefing) des im Planspiel Erlebten verweist. Das Debriefing gilt als zentraler Bestandteil von Planspielen (Peters & Vissers, 2004; Petranek et al., 1992; Schwägele et al., 2021). Durch das Debriefing, die Reflexion der im Spiel gemachten Erfahrung, unterscheiden sich Planspiele auch wesentlich von anderen Formen des spielerischen Lernens, wie den Serious Games oder Lernspielen, die nicht notwendigerweise angeleitet reflektiert werden. Für den Einsatz von Planspielen in der Hochschullehre ergänzt Schwägele: „Ziel des professionellen Einsatzes ist die Ermöglichung von Lernen.“ (Schwägele, 2015)

Entsprechend der bisherigen theoretischen Herleitung meint der Begriff Planspiel im Rahmen dieser Arbeit:

- Ein Spiel (Game), bestehend aus Actors, Rules and Resources, in dem Spielende eine temporäre soziale Organisation bilden (Klabbers)
- Das Spiel enthält der Wirklichkeit angenäherte Modelle (verschiedene Abstraktionsgrade möglich) (Kriz)
- Die im Spiel erlebten wirklichkeitsrelevanten Aus- und Folgewirkungen werden reflektiert (Debriefing) (Kriz)
- Ziel ist die Ermöglichung von Lernen (Schwägele)

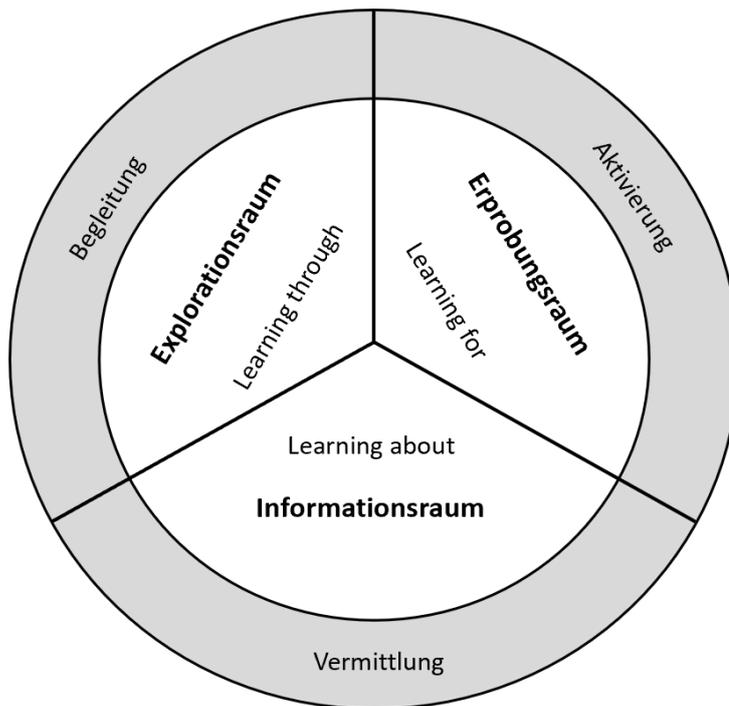
3.2 Lerntheoretische und didaktische Einordnung von Planspielen

Mit der Formulierung „Ermöglichung von Lernen“ ist bereits angedeutet, mit welcher didaktischen Haltung Planspiele (nicht nur bei Schwägele) besprochen werden. Sie legt nahe, dass Planspiele einen Rahmen bieten, in dem gelernt werden kann, nicht notwendigerweise muss. Lernende stehen als Akteure des Lernprozesses im Mittelpunkt. Planspiele werden als „constructivist learning environment“ beschrieben (Costin et al., 2018, S. 137) und vielfach explizit mit problemorientiertem Lernen in Verbindung gebracht (Costin et al., 2018; W. C. Kriz, 2010; Mandl, 2004). In anderen Texten wird für die theoretische Einordnung des Lernens mit Planspielen eher auf erfahrungsorientiertes Lernen (experiential learning) (Geithner & Menzel, 2016; Taylor et al., 2012) referenziert. Diese unterschiedlichen lerntheoretischen Verortungen setzen zwar unterschiedliche Schwerpunkte, haben aber dennoch vieles gemeinsam. Sie alle argumentieren aus einer eher konstruktivistischen Perspektive. Durch eigene Aktivität und reflektierte Erfahrung eignen sich Lernende Wissen und Kompetenzen an. Zur Einordnung von Planspielen in den Kontext der Hochschullehre greife ich das Modell der akademischen Lehre nach Reinmann (2016, S. 236) auf. Das Modell (Abbildung 2) beschreibt die Gestaltung von Hochschullehre anhand von drei metaphorischen Räumen, in denen unterschiedlich gelernt und entsprechend unterschiedlich didaktisch begleitet wird:

- Im Informationsraum steht die Vermittlung von bestehendem Wissen im Vordergrund. Studierende machen sich kundig, lernen bereits vorhandenes Wissen vorwiegend rezeptiv. Der Informationsraum ist „learning about“.
- Im Erprobungsraum können sich Studierende anhand von edukativen Fällen/Projekten/Problemen selbst erproben. Studierende werden aktiviert und wenden vorhandenes Wissen praktisch an, sie sind aktiv und produzieren eigene Resultate. Der Erprobungsraum ist „learning for“.
- Im Explorationsraum nehmen Studierende aktiv an echten (Forschungs-)Projekten teil und produzieren real bedeutsame Resultate. Sie werden pädagogisch/beratend begleitet. Der Explorationsraum ist „learning through“.

Die Räume stehen nach Reinmann in der Praxis nicht streng getrennt nebeneinander, sondern kennzeichnen sich eher durch die Akzentuierung auf einen Schwerpunkt.

Abbildung 2: Modell der akademischen Lehre nach Reinmann, 2016 (eigene Darstellung)



Im referierten Artikel bezieht sich Reinmann auf Forschungslernen. Im Informationsraum geschieht „Learning about Research“ (z.B. theoretische Einführung in Wissenschaftliches Arbeiten). Im Erprobungsraum („Learning for Research“) arbeiten Studierende mit bereitgestellten Forschungsdaten und erproben eine bestimmte Auswertungsmethode. Im Explorationsraum („Learning through Research“) werden Studierende in ein echtes Forschungsprojekt involviert und führen beispielsweise Interviews. Das Konzept lässt sich meines Erachtens jedoch auch auf andere Bereiche übertragen, z.B. Learning about Economy, Learning for Economy, Learning through Economy. Alle oben skizzierten lerntheoretischen Verortungen des Lernens mit Planspielen finden sich in dem wieder, was Reinmann den Erprobungsraum der akademischen Lehre nennt: Lernende sind aktiv und wenden Wissen an, sie treffen Entscheidungen, lösen Probleme und produzieren Ergebnisse. Ein Wirtschaftsplanspiel ist Learning for Business, ein Politikplanspiel ist Learning for Politics. Learning about Business

oder Learning about Politics hieße, sich vor allem rezeptiv mit Grundbegriffen und Theorien eines Fachs zu beschäftigen (z.B. in Vorlesungen, Seminaren oder im Selbststudium). Learning through Business oder through Politics wäre es, wenn sich Studierende im Rahmen einer Lehrveranstaltung tatsächlich wirtschaftlich oder politisch engagieren würden, also den geschützten Rahmen eines Spiels, der Simulation oder eines Projektes verlassen würden. Für den Bereich Wirtschaft findet dieses Konzept beispielsweise in den Team Academics Anwendung. In dem aus Finnland stammenden Modell gründen Studierende im Rahmen ihres Studiums ein Unternehmen und lernen Management vorwiegend anhand der eigenen Praxis (Tosey et al., 2015)⁴.

Das Bildungswesen ist keineswegs das einzige Einsatzgebiet für Planspiele. Planspiele werden auch in eher experimenteller Weise eingesetzt. So nutzt Freese (2019) ein Planspiel zum Flughafenmanagement (und die Spielenden als Probanden) für ihre Untersuchung zum Einfluss von Emotionen auf die Teamarbeit. Kawalle (2017) diskutiert und erprobt den Einsatz von politischen Planspielen zur empirischen Überprüfung von politikwissenschaftlichen Theorien. Im Versuch einer Systematisierung der Einsatzzwecke von Planspielen nennen Kriz und Manahl (2018) neben „*Education and Training*“ zwei weitere: „*simulation and game-based policy interventions*“ und „*gaming simulation as a core method of design of complex socio-technical systems*“. Im Kontext dieser Arbeit werden Planspiele jedoch ausschließlich als Lehr-Lernmethode diskutiert und untersucht.

3.3 Warum Planspiele? Versuch einer Begründung

Planspiele sind in großer Anzahl vorhanden und decken ein breites Spektrum an unterschiedlichen Themen ab. Die Planspieldatenbank der Bundeszentrale für Politische Bildung umfasst derzeit 250 Planspiele. Berücksichtigt sind hier nur deutschsprachige Planspiele für die politische Bildung, während Themen wie Wirtschaft oder Technik ausgeklammert bleiben (<https://www.bpb.de/lernen/angebote/planspiele/datenbank-planspiele/>). Das Bundesinstitut für Berufsbildung zählt über 500 Planspiele und Serious Games mit überwiegend wirtschaftlichen Themen auf (Blötz, 2015). Neben den großen

⁴ Seit 2018 kann nach diesem Modell an der Hochschule Bremerhaven der Studiengang Gründung, Innovation, Führung studiert werden: <https://www.hs-bremerhaven.de/Gruendung-Innovation-Fuehrung>.

Einsatzfeldern Politik und Wirtschaft werden auch in anderen Fachbereichen Planspiele entwickelt und eingesetzt. Neben IT-Planspielen (Beranic & Hericko, 2019) finden sich Planspiele im Gesundheitswesen⁵ (Ney et al., 2014), in der Theologie (Riegger, 2023) und vermehrt auch in der Lehrkräftebildung (Kadel et al., 2023). Daraus ergibt sich die Frage, mit welcher Motivation Planspiele in einer solchen Breite entwickelt und in der Lehre eingesetzt werden. Aus welchem Grund erfreuen sie sich einer solchen Beliebtheit? Neben jeweils sehr spezifischen Antworten (z.B. Erlernen des in einem bestimmten Planspiel abgebildeten Themas) können auch übergreifende Antworten jenseits von konkreten Fachthemen gegeben werden. Ohne einen Anspruch auf Vollständigkeit erheben zu wollen, lassen sich zusammenfassend insbesondere drei übergreifende Gründe nennen, warum Planspiele in der Lehre eingesetzt werden: die Multidimensionalität des Lernens, die Tiefe des Lernens und das Erleben von systemischen Zusammenhängen.

Planspiele sind „multidimensionale Lehr-/Lernarrangements“: Sie ermöglichen es, „Fachwissen zu erwerben, sozial-kommunikative Fertigkeiten einzuüben und methodisch-strategische Kompetenzen zu erlernen.“ (Engartner & Meißner, 2017). Engartner und Meißner betonen, dass mit Planspielen gleichzeitig mehrere Lernmöglichkeiten eröffnet werden. Sachthemen sind stets eingebunden in einen konkreten Anwendungsfall. Wirtschaftliche Kennzahlen werden gelernt und zugleich zur Entscheidungsfindung im Planspiel angewendet. Formelle demokratische Entscheidungsprozesse werden gelernt und zugleich im Planspiel durchlaufen. Der Anwendungsfall ist zudem stets in einen sozialen Kontext eingebunden. Lernende erklären sich die wirtschaftliche Kennzahl beispielsweise gegenseitig und diskutieren, wie sie anhand dieser zu einer Entscheidung kommen. In politischen Planspielen nehmen Teilnehmende an einem simulierten Wahlkampf teil und diskutieren aus der Perspektive ihrer jeweiligen Rolle. Das Erlernen bestimmter Themen findet nicht isoliert statt, es ist integriert in ein sinnvolles Netz aus sozialen und methodischen Anforderungen, die Lernende gleichzeitig bearbeiten müssen. Das Lernen wird in Planspielen so zu einer mehrdimensionalen, ganzheitlichen Erfahrung.

⁵ Im Gesundheitswesen wird auch häufig mit anderen simulativen Methoden (z.B. Puppen, die bestimmte Symptome simulieren) gearbeitet (Al-Eliq (2010)). Der Übergang zu Planspielen kann fließend sein.

Um Lernen und Lernziele differenziert beschreiben zu können werden häufig Taxonomien verwendet. Die weiterentwickelte Taxonomie nach Bloom (Krathwohl, 2002) unterscheidet beispielsweise verschiedene Dimensionen von Wissen (Factual Knowledge, Conceptual Knowledge, Procedural Knowledge, Metacognitive Knowledge) einerseits und den kognitiven Prozess, in dem dieses Wissen angewendet wird andererseits (Remember, Understand, Apply, Analyze, Evaluate, Create). Da Planspiele eine sehr handlungsorientierte Lehr-Lernmethode sind, kann angenommen werden, dass mit ihnen, im Sinne der Taxonomie, tieferliegende Ebenen des Lernens erreicht werden können. Planspiele setzen voraus, dass Inhalte nicht nur erinnert und verstanden werden, sondern, dass aufgrund der verarbeiteten Informationen Entscheidungen getroffen werden. Die Inhalte müssen also angewendet, analysiert und evaluiert werden. Planspiele machen es demnach erforderlich Inhalte auf tieferliegenden Ebenen zu bearbeiten.

In Planspielen werden Lernende nicht mit einzelnen Problemen, sondern mit ganzen Systemen und ihren vielseitigen kausalen und funktionalen Wechselwirkungen konfrontiert; sie sind somit aufgefordert systemisch zu denken (Geuting, 2000). „A system is an interconnected set of elements that is coherently organized in a way that achieves something. (...) a system must consist of three kinds of things: elements, interconnections, and a function or purpose.“ (Meadows, 2009, S. 11). Was sich in der Theorie sehr überschaubar anhört, kann in der Realität schwer zu durchschauen sein: Welche Elemente gehören zum System, wie sind die Verbindungen und welchen Zweck erfüllt das System? Systeme können nach Monat und Gannon (2015, S. 17) mit linearen Denkmustern nicht adäquat erfasst werden. Es brauche holistisches Denken, das multiple räumliche und zeitliche Einflüsse zugleich einbeziehe. Nach Duke und Geurts (2004, S. 271) sollen Planspiele gerade systemische Zusammenhänge abbilden: „the intended end product of the design progress is an operating model of a real-life system. In the game model, people in different but interrelated roles create, at least partly, the dynamic behavior of the model.“ In Planspielen erleben Lernende, wie Systemelemente in zeitlicher und räumlicher Weise interagieren und können zudem ausloten, mit welchen Handlungsmöglichkeiten sie sich in einem (politischen oder ökonomischen) System „bewegen“ können. Hierzu zählen auch „realitätsnahe Probehandlungen“ (Ausprobieren von Verhaltensmustern, Strategien) und der „Perspektivwechsel“ durch die Übernahme fremder Rollen

(Engartner & Meßner, 2017). Als dritter allgemeiner Grund für die Nutzung von Planspielen in Lehrkontexten kann gefolgert werden: Indem Planspiele systemische Zusammenhänge simulieren, eignen sie sich zum handlungsorientierten (ausprobieren, Rollenübernahme) Erlernen komplexer systemischer Zusammenhänge.

3.4 Typologisierung von Planspielen

Da es eine beinahe unüberschaubare Menge an unterschiedlichen Planspielen gibt, die sehr unterschiedliche Spielziele verfolgen und mit sehr unterschiedlichen Spielmechanismen arbeiten, scheint eine Typologisierung sinnvoll.

Eine häufige Einteilung von Planspielen findet sich entlang von Fachgrenzen bzw. anhand von Themen, die mit den Spielen adressiert werden. So finden sich Veröffentlichungen, die sich speziell politischen Planspielen widmen (Petrik & Rappenglück, 2017) und andere, die sich eher mit Wirtschaftsplanspielen beschäftigen (Blötz, 2003). Des Weiteren können Planspiele, unabhängig von ihrem Thema, anhand ihrer äußeren Form in haptische Planspiele, computergestützte Planspiele und Mischformen hiervon eingeteilt werden (Zürn et al., 2023).

Eine relativ abstrakte Möglichkeit der Typologie findet sich bei Klabbers (2018). Hier werden Spiele anhand ihrer inneren Logik, ihres Spielprinzips kategorisiert. Anhand von Values unterscheidet Klabbers Spiele in Type I and Type II Games (Tabelle 2). Type I Games sind durch Einigkeit unter den Spielenden bezüglich der im Spiel zu erstrebenden Werte gekennzeichnet (z.B. möglichst hoher Aktienkurs). Es entsteht Wettkampf/Konkurrenz, da sich die Mitspielenden um den zu erreichenden Wert streiten. Type II Games sind dadurch gekennzeichnet, dass zwischen den Spielenden Dissens über die zu erstrebenden Werte herrscht. Spielende stehen nicht in Konkurrenz, sondern in Konflikt miteinander. Sie streiten nicht um das Erreichen bestimmter Ziele, sondern um die „richtigen“ Werte und Ziele. Durch die Einführung einer weiteren Achse (Knowledge) differenziert Klabbers weiter: Teilnehmende können sich darauf verlassen, dass das Spiel Wissen zur Problemlösung bereitstellt einerseits und Spielende sind völlig ungewiss, ob das Spiel Wissen zur Problemlösung bereithält andererseits. Anhand dieser Unterscheidungsachsen ergeben sich vier mögliche Spieltypen. In *Manageable knowledge problems* herrscht Einigkeit über das zu erreichende Ziel (Konkurrenz) und Spielende können sich darauf verlassen, dass das Spiel

sämtliche relevante Informationen zur Verfügung stellt. In Spielen des Typs *Tricky ethical problems* finden Spielende ebenfalls alle relevanten Informationen vor, es herrscht jedoch Dissens über die zu erreichenden Ziele unter den Spielenden (Konflikt). Bei Spielen des Typs *(In-)tractable knowledge problems* sind sich Spielende einig über die zu erstrebenden Ziele, müssen sich die Informationen im Laufe des Spiels jedoch selbst erarbeiten. *Wicked governance problems* sind auf beiden Achsen durch ein hohes Maß an Unsicherheit gekennzeichnet: Weder können sich Spielende auf bereitgestellte Informationen verlassen, noch herrscht Einigkeit über gemeinsame Werte.

Tabelle 2: Typologisierung von Planspielen nach Klabbers (2018)

			Knowledge	
			<i>Certainty about knowledge to be, or to become available</i>	<i>Uncertainty about knowledge</i>
<i>Type-I Games</i>	Values	<i>Consensus (competitive goals)</i>	Manageable knowledge problems	(In-)tractable knowledge problems
<i>Type-II Games</i>		<i>Dissension (conflicting goals)</i>	Tricky ethical problems	Wicked governance problems

Tabelle 3 gibt einen Überblick über die im empirischen Teil dieser Arbeit verwendeten Planspiele⁶ und ordnet sie entsprechend der skizzierten Typisierungen ein. Die Datenerhebung erfolgte am Fachbereich Wirtschaft der Dualen Hochschule Baden-Württemberg Stuttgart. Somit lassen sich die meisten Planspiele dem Themenkomplex Betriebswirtschaftslehre/Wirtschaftswissenschaften zuordnen. In typischen Wirtschaftsplanspielen übernehmen Studierende zumeist eine Managementrolle und treffen Unternehmensentscheidungen. Über die zu erreichenden Ziele und Werte herrscht hierbei Konsens. Der Spielerfolg zeigt sich im Erreichen bestimmter KPIs (Aktienkurs, Umsatz, Gewinn oder ähnlich), weshalb ein Wettbewerb zwischen den Teams entstehen kann. Es sind demnach Type-I Spiele. Neben Unternehmensplanspielen stehen immerhin 5 Beispiele für Type-II Spiele. Mit

⁶ Weitere Informationen und Verlinkungen zu den Herstellern der Planspiele finden sich unter: <https://zms.dhbw-stuttgart.de/das-zms/unsere-planspiele/>.

ihnen werden Themen wie Nachhaltigkeit, interdisziplinäre Zusammenarbeit oder Organisationsentwicklung thematisiert. Hierbei ist unter den Spielenden (je nach Rolle) nicht klar, welche Ziele/Werte erreicht werden sollen und es entstehen Konflikte. Schwieriger als die Einordnung in Type-I und Type-II Spiele fällt die Einteilung über das Vorhandensein von Wissen, da in jedem Spiel ein gewisses Maß an Wissen über das Spiel und seine Regeln vorhanden sein muss. Als pragmatisches und relativ gut operationalisierbares Unterscheidungsmerkmal dient daher das Vorhandensein eines den Teilnehmenden zur Verfügung stehenden Handbuchs, in dem prinzipiell alle relevanten Informationen zugänglich sind. Ist das Spiel jedoch so konzipiert, dass Teilnehmende relevante Informationen bei der Spielleitung erfragen bzw. von anderen Mitspielenden beziehen müssen, ohne dass sie explizit darauf hingewiesen werden.

Tabelle 3: Planspiele nach Typisierung

	Name	Thema der Simulation	Wissen vorhanden	Wissen nicht vorhanden
Type-I Games	coludo-Industry	Management, Accounting	x	
	TOPSIM General Management	Management	x	
	TOPSIM Hospital Management	Management, Gesundheitswesen	x	
	BTI Store Manager	Management, Handel	x	
	BTI Hospital Akut	Management, Gesundheitswesen	x	
	TOPSIM Logistics	Management, Logistik	x	
	riva SysTeamsChange	Change-Management		x
	TOPSIM Universal Banking	Management, Finanzierung	x	
	MSD-Game	Management	x	
	TOPSIM Insurance	Management, Versicherung	x	
	Ökonomikus Insurance	Management, Versicherung	x	
	AMGD	Agiles Projektmanagement		
	TOPSIM BD Fitness	Management	x	
	BTI Factory	Management, Accounting	x	
	BTI Global Strategy	Management	x	
	TOPSIM Easy Management	Management	x	
	CAPSIM-Core	Management	x	
	TOPSIM Going Global	Management, globale Expansion	x	
	TOPSIM Manufacturing Management	Management, Herstellung	x	
	TOPSIM Scale Up	Management, Disruptive Märkte	x	
	HR-Management	Management, Personal	x	
	TOPSIM Applied Economics	Volkswirtschaftslehre,	x	
	BTI Logistic World Competition	Management	x	
	Inchainge - The Fresh Connection	Management, Logistik	x	
	TOPSIM Immo Makler	Management, Immobilien	x	

Type-II Games	Fountains	Ressourcen, Kooperation	x
	Fishing Game	Nachhaltigkeit, Ressourcen	x
	Grid	Management, Organisationsentwicklung	x
	Mission Team Development	Management, Kooperation, Organisationsentwicklung	x
	ProCom Medieval Market	Management, Interprofessionalität	x

4 Evaluation planspielbasierter Lehrveranstaltungen: Beschreibung des Fragebogens

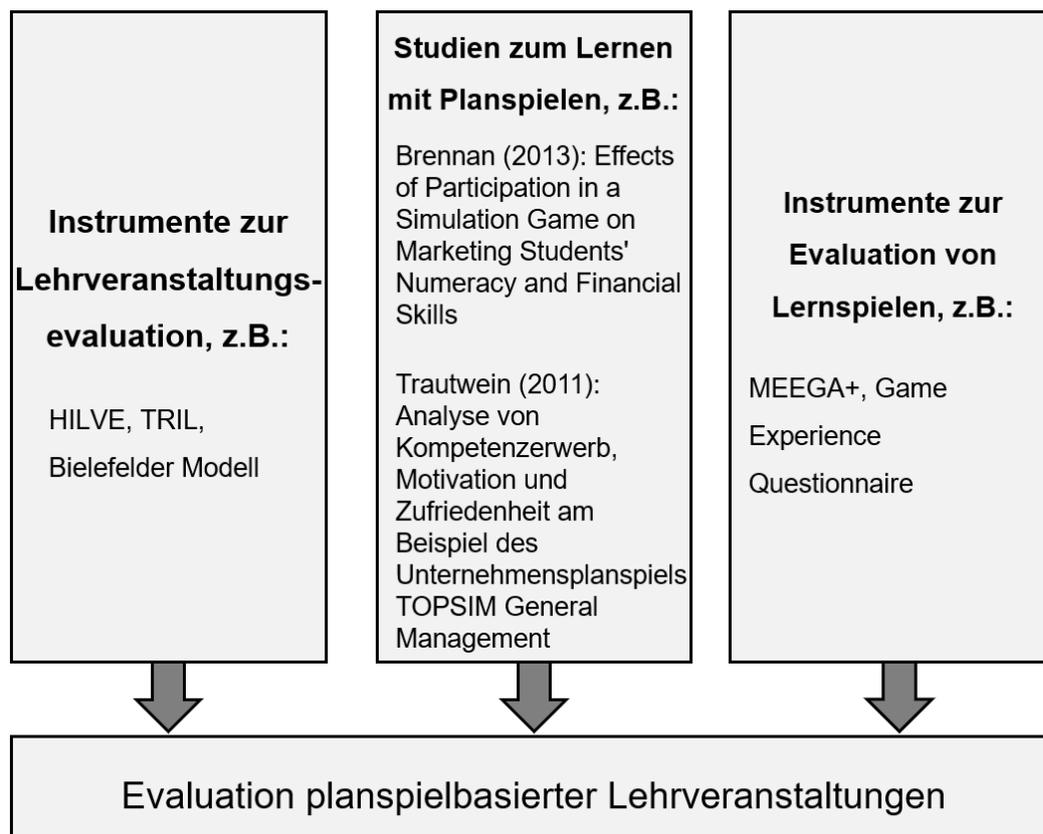
Der empirische Teil dieser Arbeit widmet sich der Evaluation planspielbasierter Lehrveranstaltungen, wofür ein geeigneter Fragebogen notwendig ist. Die Fragebogenentwicklung erfolgte am ZMS. Die Konzeption und erste Daten zur Fragebogenentwicklung wurden bereits präsentiert und veröffentlicht (F. Trautwein & Alf, 2023). Im Folgenden wird skizziert an welche Vorarbeiten bei der Fragebogenentwicklung angeknüpft wurde und in welchen Schritten der ZMS-Fragebogen entwickelt wurde (wobei immer wieder auf die genannte Veröffentlichung verwiesen wird). In den Kapiteln 4.3 und 4.4. wird die Fragebogenstruktur anhand jüngerer Daten auf Reliabilität und Validität getestet.

4.1 Anknüpfung an unterschiedliche Forschungsstränge

Für die Entwicklung eines solchen Instruments kann an drei unterschiedlichen Forschungssträngen angeknüpft werden (siehe Abbildung 3). Erstens kann an Konzepte der **Lehrveranstaltungsevaluation** angeknüpft werden, wie sie in Kapitel 2.4 beschrieben wurden. Diese Fragebögen adressieren in der Regel klassische Lehrveranstaltungen (z.B. Vorlesung, Seminar, Lehrgespräch). Damit geht einher, dass sie Aspekte, die in der Lehre mit Planspielen relevant sind, nicht berücksichtigen. Für die Entwicklung eines geeigneten Fragebogens zur Evaluation planspielbasierter Lehrveranstaltungen können Inventare zur

klassischen Lehrveranstaltungsevaluation eine gute Orientierung bieten. Es müssen jedoch noch weitere Elemente berücksichtigt werden, wie z.B. die Qualität des Planspiels oder die Teamarbeit. Zweitens können Studien betrachtet werden, die **Lernprozesse mit Planspielen** untersuchen. Hier findet sich ein breites Feld an Studien sehr unterschiedlicher Qualität (Susan Zeiner-Fink et al., 2023). Diese Studien konzentrieren sich zumeist auf ein konkretes Planspiel: Es wird erforscht, inwiefern ein bestimmtes Lernziel (z.B. Financial Skills, vgl. Abbildung 3) mit einem bestimmten Planspiel gefördert werden kann. Da diese Studien stets konkrete Planspiele mit ihren spezifischen Mechanismen und Lernzielen vor Augen haben, sind die verwendeten Items für die Entwicklung eines allgemeinen Instruments zur Evaluation planspielbasierter Lehre meist zu spezifisch. Die Entwicklung eines allgemeineren Fragebogens kann sich daran orientieren, muss jedoch in vielen Formulierungen abstrakter bleiben, um für eine Vielzahl an Planspielen Anwendung finden zu können. Als dritte Inspirationsquelle können Forschungsinstrumente zur **Evaluation von Lernspielen** herangezogen werden. Hierzu zählen beispielsweise der Game Experience Questionnaire (Ijsselsteijn et al., 2013) und das Model for the Evaluation of Educational Games (MEEGA+) (Petri et al., 2017). Beide Instrumente legen den Fokus auf bestimmte **Merkmale von (Lern-)Spielen**. Diese Fragebögen sind für die Entwicklung eines Fragebogens zur planspielbasierten Lehrveranstaltungsevaluation zu beachten, weil sie über ein ähnliches Abstraktionsniveau verfügen. Die Items sind so formuliert, dass sie für möglichst viele unterschiedliche Spiele relevant sind. Sie haben jedoch Lernspiele vor Augen, die ohne Betreuung und Anleitung gespielt werden, haben also nicht die Einbindung in eine formelle Lehrveranstaltung im Blick.

Abbildung 3: Anknüpfung an Forschungsstränge



Für die Fragebogenentwicklung wurden alle drei Perspektiven in Betracht gezogen. Sie finden sich in unterschiedlicher Stärke im finalen Fragebogen wieder. Dieser wird im folgenden Kapitel vorgestellt.

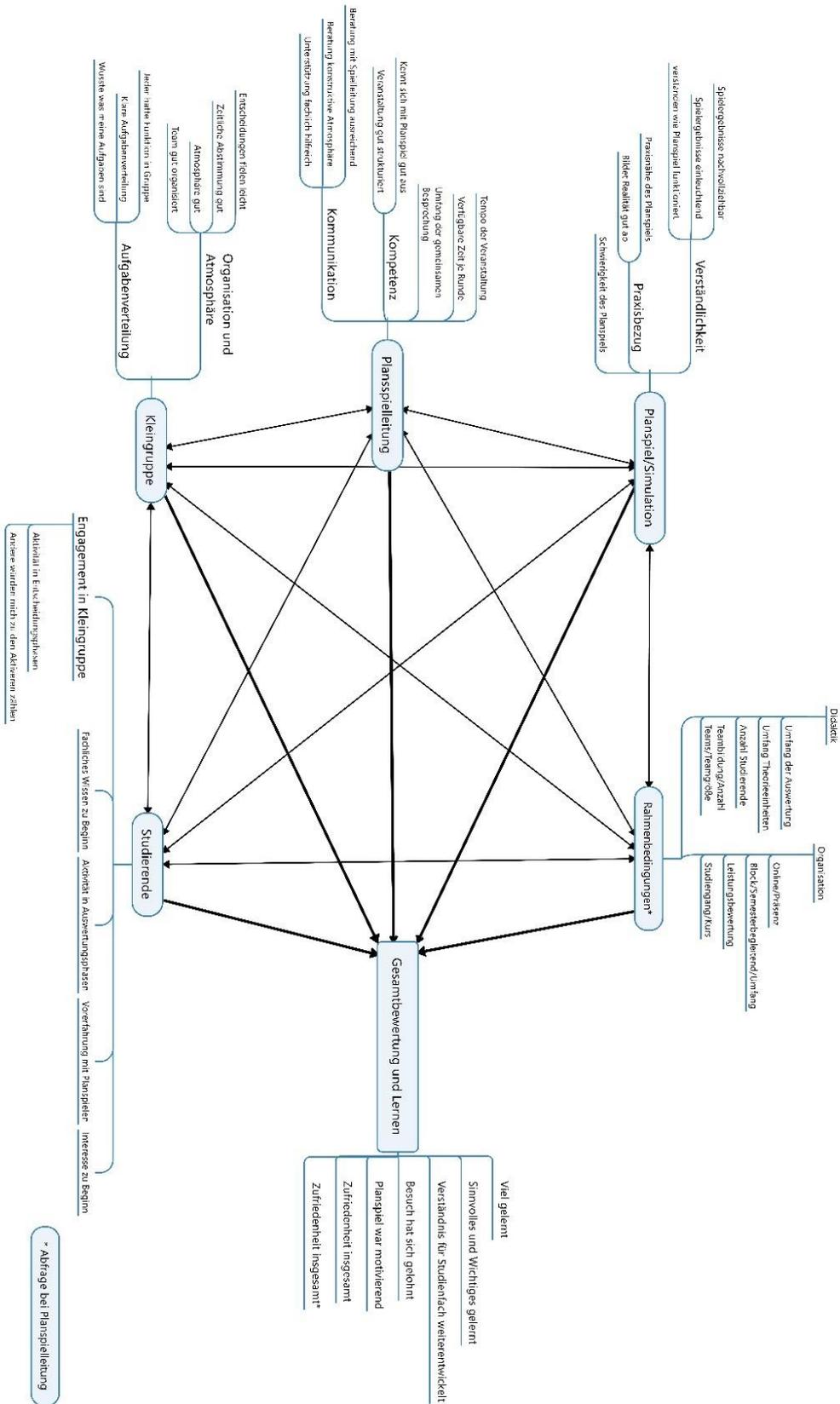
4.2 Der ZMS-Fragebogen

Ein Ziel dieser Arbeit ist es, planspielbasierte Lehrveranstaltungen in hoher Zahl zu evaluieren. Hierzu wird ein Evaluationsinstrument benötigt, das unkompliziert und ökonomisch einsetzbar ist. Aufwendige Forschungsmethoden mit Kontroll- und/oder Pre-Post-Designs kommen zu diesem Zweck kaum in Frage. Ziel war es daher einen Fragebogen zu konzipieren, der eine Selbsteinschätzung zum Lernen mit Planspielen erhebt, sowie relevante Einflussfaktoren. Der Bogen sollte am Ende der Lehrveranstaltungen in ca. 5-10 Minuten auszufüllen sein, wie dies in der Lehrveranstaltungsevaluation häufig der Fall ist. Inspirationsquellen hierfür kamen, wie beschrieben, aus drei Richtungen: Lehrveranstaltungsevaluation, Forschung zu spezifischen Planspielen, Forschungsinstrumente zu Game-Design.

Für die Entwicklung des Fragebogens für Studierende wurde zunächst mit 79 Items gearbeitet. Mit diesen wurden (F. Trautwein & Alf, 2023) Post-Befragungen

in 10 Veranstaltungen durchgeführt und 182 Datensätze erhoben. Die Items wurden theoriegeleitet den folgenden Dimensionen zugeordnet (siehe Anhang): Gesamtbewertung, Lernen, Rahmenbedingungen, Planspiel, Planspielleitung, Debriefing/Interaktion, Teamarbeit, Studierende. Die angenommene Faktorenstruktur wurde mit Reliabilitätsanalysen und Explorativen Faktorenanalysen überprüft. Aufgrund der statistischen Ergebnisse wurde die Faktorenstruktur angepasst und ein finaler Fragebogen entwickelt (F. Trautwein & Alf, 2023). Für die Dimensionen Gesamtbewertung und Lernen zeigt sich eine deutliche Ladung auf einen Faktor, sodass diese zu einer Dimension Lernen und Gesamtbewertung zusammengefasst wurden (ebd. S. 9). Zudem konnten die Dimensionen Planspielleitung einerseits und Debriefing/Interaktion andererseits nicht diskriminant voneinander dargestellt werden. Auch hier zeigen sich deutliche Ladungen auf einen Faktor, sodass Items aus den ursprünglich angenommenen zwei Dimensionen zu einer Dimension Planspielleitung bzw. Facilitation zusammengefasst wurden (ebd.). Alle anderen Dimensionen erwiesen sich als stabil. Des Weiteren wurden Items mit geringer Ladung und/oder hoher Querladung aussortiert. Ein kürzerer Fragebogen wird auch an Lehrende ausgegeben. Sie werden zu ihrer Vorerfahrung mit Planspielen befragt und geben ebenfalls eine kurze Einschätzung zum Gelingen der Veranstaltung und zur Mitarbeit der Studierenden ab. Im Zentrum des Fragebogens für Lehrende stehen jedoch die Rahmenbedingungen, die vorwiegend mit nominalen Fragen abgedeckt werden. Abgefragt wird beispielsweise, ob mit dem Planspiel eine Prüfungsleistung verbunden ist, ob das Planspiel in Präsenz oder Onlinelehre stattfindet oder inwiefern weitere Theorieeinheiten mit dem Planspiel verbunden werden. Abbildung 4 veranschaulicht die in der Evaluation berücksichtigten Dimensionen und skizziert die angenommene interdependente Struktur.

Abbildung 4: Evaluationsmodell



Die folgenden Analysen beziehen sich auf den finalen Fragebogen für Studierende, wie er für die empirischen Studien in Kapitel 5 verwendet wurde, und diskutieren seine Reliabilität und Validität.

Die Durchführung der folgenden statistischen Berechnungen erfolgte mit Jamovi, einer auf der Programmiersprache R basierenden grafischen Oberfläche (The jamovi project, 2023). Die statistischen Analysen zur Reliabilität und Validität des entwickelten Fragebogens basieren zumeist auf 3067 Datensätzen, die zwischen Mai 2021 und April 2023 am ZMS der DHBW Stuttgart erhoben wurden. Nur der zuletzt verfasste Artikel (Kapitel 5.4) umfasst genau diese Anzahl von Fällen. Früher veröffentlichte Artikel arbeiten mit kleineren Fallzahlen, wobei sich Reliabilität und Validität auch mit kleineren Fallzahlen als stabil erwiesen haben.

4.3 Explorative Faktorenstruktur

Explorative Faktorenanalysen (EFA) werden verwendet, um Strukturen in Daten zu erkunden (Exploration); häufig dann, wenn noch keine theoretischen Annahmen bestehen oder eben zur Entwicklung von Fragebögen (Backhaus et al., 2018, S. 366; Eid et al., 2013, S. 893). Eid et al. beschreiben EFA in drei Schritten: Bestimmung der Anzahl der Faktoren, Anfangslösung und Rotation. Für die Bestimmung der Faktoren schlagen Eid et al. vor, zunächst von einem Faktor auszugehen und die Anzahl schrittweise zu erhöhen, bis ein zufriedenstellendes Modell entsteht (S. 895). Jamovi bietet zudem zwei Verfahren zur computergestützten Errechnung der Faktorenanzahl: Eigenwerte und Parallelanalyse. Der Eigenwert beschreibt, welche Bedeutung ein Faktor im Vergleich zu anderen Faktoren des Modells einnimmt, wie viel Varianz in den Daten durch einen Faktor erklärt wird. Üblicherweise werden Faktoren mit Eigenwerten von 1 und höher als relevante Faktoren interpretiert (Kaiser-Guttman-Kriterium) (Bortz & Schuster, 2010, S. 415). Die Parallelanalyse nutzt ein rechnerisch sehr aufwändiges Verfahren (das daher erst nach und nach Eingang in Statistiksoftware findet). Hierbei werden die Eigenwerte der Faktorenanalyse mit den Eigenwerten von Zufallsdaten verglichen. Beibehalten werden diejenigen Faktoren, deren Eigenwert größer ist als der Eigenwert in den Zufallsdaten (Bortz & Schuster, 2010, S. 416).

Unter Nutzung der Parallelanalyse ergeben sich sieben Faktoren, die sehr stark das theoretisch angenommene Konstrukt widerspiegeln (Tabelle 4). Die sechs Items zu Lernen und Gesamtbewertung laden allesamt auf den Faktor 1. Die

Dozierendenvariablen laden allesamt stark auf Faktor 2. Bemerkenswert ist, dass sich die Team-Items auf zwei Faktoren verteilen: Einen Faktor (Faktor 3), in dem sich Variablen gruppieren, die allesamt beschreiben wie „gut“ die Teamarbeit war oder wie „leicht“ Entscheidungen getroffen werden konnten. Es geht hierbei eher um atmosphärische Aspekte der Teamarbeit. Unter Faktor 5 gruppieren sich hingegen Variablen, die die Aufgabenverteilung und die Funktionen innerhalb des Teams beschreiben. Faktor 4 fasst Variablen zusammen, die beschreiben, wie verständlich, einleuchtend und nachvollziehbar das Planspiel für die Studierenden war. Das Item P3 lädt mit knapp unter 0.5 eher schwach auf den Faktor, wobei keine nennenswerten Querladung vorhanden sind. Auf Faktor 6 laden die drei Items, die erfragen, wie die Studierenden ihre eigene Mitarbeit, ihr eigenes Engagement bewerten. Mit 0.426 lädt das Item S1 eher schwach auf den Faktor, wobei auch hier keine Querladungen vorhanden sind. Auf Faktor 7 laden zwei weitere Planspiel-Items. Sie beschreiben, wie realitäts- und praxisrelevant das Planspiel eingeschätzt wird. Beide Items von Faktor 7 laden auch mit 0.382 und 0.359 auf Faktor eins. Diese Querladung ist nennenswert, jedoch deutlich (>0.3) im Abstand zur Querladung auf Faktor 7.

Tabelle 4: Explorative Faktorenanalyse (N=3067)

		Faktoren						
		1	2	3	4	5	6	7
G_1	Alles in allem hat sich der Besuch dieser Veranstaltung für mich gelohnt.	0.831						
L_2	Ich habe in dieser Veranstaltung etwas Sinnvolles und Wichtiges gelernt.	0.816						
G_2	Das Planspiel war für mich sehr motivierend.	0.786						
L_1	Ich habe beim Planspiel viel gelernt.	0.774						
L_3	Mein Verständnis für mein Studienfach hat sich durch die Veranstaltung weiterentwickelt.	0.731						
G_3	Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?	0.692	0.408					
Inter. Anfang	Mein Interesse am Planspiel war zu Beginn der Veranstaltung...	0.309						
D_4	Die Beratungen mit der Planspielleitung fanden in einer konstruktiven Atmosphäre statt.		0.797					
D_5	Die Unterstützung der Planspielleitung war fachlich hilfreich.		0.785					
D_3	Es bestand in ausreichendem Maß die Möglichkeit, sich mit der Planspielleitung zu beraten.		0.712					
D_2	Die Veranstaltung war gut strukturiert.		0.689					
D_1	Die Planspielleitung kennt sich mit dem Planspiel gut aus.		0.669					
T_2	Die zeitliche Abstimmung in der Gruppe hat gut funktioniert.			0.764				
T_1	Es fiel uns leicht, in der Gruppe Entscheidungen zu treffen.			0.736				
T_4	Ich würde sagen, dass unser Team sehr gut organisiert war.			0.672		0.332		
T_3	Die Atmosphäre in der Gruppe war gut.			0.632				
P_2	Die Spielergebnisse sind einleuchtend.				0.833			
P_1	Die Spielergebnisse lassen sich gut nachvollziehen.				0.822			
P_3	Ich habe verstanden, wie das Planspiel funktioniert.				0.477			
T_6	Es gab eine klare Aufgabenverteilung in der Gruppe.					0.863		
T_5	Jeder hatte eine Funktion in der Gruppe.			0.333		0.684		
T_7	Ich wusste, was meine Aufgaben sind.					0.590		
S_3	Wenn andere meine Rolle betrachten, würden sie mich zu den aktiveren Teilnehmenden zählen.						0.902	
S_2	In den Entscheidungsphasen (Arbeit in der Kleingruppe) habe ich mich aktiv beteiligt.						0.689	
S_1	In den Auswertungsphasen (gemeinsame Besprechung der Spielergebnisse im Plenum) habe ich mich aktiv beteiligt.						0.426	
P_4	Das Planspiel ist praxisnah.	0.382						0.772
P_5	Das Planspiel bildet die Realität gut ab.	0.359						0.670
Anmerkung: „Maximum likelihood“-Extraktion wurde zusammen mit einer „Varimax“-Rotation verwendet.								

4.4 Reliabilität und Validität

Um die explorativ gefundene Faktorenstruktur bestätigen zu können, sollten die Daten hinsichtlich ihrer Reliabilität und Validität überprüft werden.

Reliabilität beantwortet die Frage nach der Genauigkeit und der daraus resultierenden Verlässlichkeit einer empirischen Messung. Die Verlässlichkeit von Messungen erhöht sich dadurch, dass ein bestimmtes Merkmal mit mehreren Instrumenten gemessen wird, die jeweils genau dieses Merkmal abfragen sollen. In der empirischen Sozialforschung werden daher mehrere Items verwendet und zu einem Konstrukt zusammengefasst, um ein Merkmal zu messen. Fraglich ist jedoch, ob die formulierten Items tatsächlich das gleiche Merkmal messen, d.h. ob Teilnehmende an der Umfrage die hypothetisch zusammengehörenden Items tendenziell gleich beantworten. Um diese Frage zu beantworten, also die Verlässlichkeit (Reliabilität) einer Skala festzustellen, wurden verschiedene statistische Verfahren entwickelt. Im Rahmen dieser Arbeit wird die Reliabilität vorwiegend über die interne Konsistenz der Skalen erhoben. Neben dem zu diesem Zweck standardmäßig verwendeten Cronbach`s Alpha (Taber, 2018) wird zusätzlich McDonald`s Omega verwendet. McDonald`s Omega kommt zu besseren Reliabilitätsabschätzungen bei Skalen, die aus wenigen Items gebildet werden (was hier teilweise der Fall ist) und in Fällen, in denen die verwendeten Items eine Skala unterschiedlich stark repräsentieren (Bortz & Döring, 2006, S. 468; Hayes & Coutts, 2020). Interpretiert werden beide Indikatoren auf die gleiche Weise. Je näher die Werte sich 1 annähern, desto eher ist die Reliabilität gegeben, wobei sehr hohe Werte ($>0,9$) für eine gewisse Redundanz in den Items sprechen (Taber, 2018). Die Alpha und Omega Werte in Tabelle 5 lassen auf eine gute Reliabilität der Daten schließen. 5 von 7 Skalen liegen sowohl für Alpha als auch für Omega zwischen 0,8 und 0,9. Die Items bilden die Skala mit hoher Übereinstimmung verlässlich ab. Die Skala Lernen und Gesamtbewertung mit Werten über 0,9 könnte aus etwas weniger Items gebildet werden. Einige Items scheinen redundante Inhalte widerzugeben. Um jedoch das Gleichgewicht in der Skala zu wahren (3 Items fragen lernspezifische Inhalte ab, 3 Items fragen nach der Gesamtbewertung) bleiben dennoch alle Items enthalten. Etwas niedrigere Werte für die interne Konsistenz und die höchste Diskrepanz zwischen Alpha und Omega zeigt die Skala Studierende Engagement. Das Item S_1, das die Mitarbeit der Studierenden im Plenum abfragt, korreliert verhältnismäßig schwach mit der Skala. Dies ist inhaltlich plausibel. Gegenüber der engagierten Mitarbeit in der

Kleingruppe (S_2) und der engagierten Mitarbeit allgemein (S_3) muss die Mitarbeit im Plenum geringer ausfallen, da im Plenum den Dozierenden eine aktivere Rolle zufällt und die Möglichkeiten zur Mitarbeit eingeschränkter sind als in der Kleingruppe. Da die Items die Skala unterschiedlich stark repräsentieren (schwächere Ladung von S_1), scheint es hier angemessen zu sein eher dem Omega-Wert zu vertrauen, da Alpha die interne Konsistenz unterschätzen könnte (Bortz & Döring, 2006, S. 468). In den durchgeführten Studien wird das Studentische Engagement teilweise ohne das Item S_1 repräsentiert, teils um den Vorgaben der Journals Rechnung zu tragen, teils, um in der Arbeit mit Strukturgleichungsmodellen nicht gegen methodisch geforderte Schwellenwerte zu verstoßen.

Tabelle 5: Faktorenstruktur, Deskriptive Statistik, Reliabilität (N=3067)

Faktor	Indikator	Ladung auf Faktor	p	Mittelwert	Std.-abw.	Cronbachs α	McDonald's ω
Planspiel Verständnis	P_1	0.899	< .001	4.84	0.94	0.83	0.85
	P_2	0.904	< .001				
	P_3	0.593	< .001				
Planspiel Praxisbezug	P_4	0.882	< .001	4.29	1.07	0.85	0.86
	P_5	0.847	< .001				
Planspielleitung (Dozierende)	D_1	0.695	< .001	5.33	0.81	0.88	0.88
	D_2	0.744	< .001				
	D_3	0.731	< .001				
	D_4	0.821	< .001				
	D_5	0.847	< .001				
Team Atmosphäre	T_1	0.731	< .001	5.13	0.83	0.83	0.83
	T_2	0.734	< .001				
	T_3	0.696	< .001				
	T_4	0.79	< .001				
Team Arbeit	T_5	0.799	< .001	4.6	1.21	0.82	0.82
	T_6	0.836	< .001				
	T_7	0.702	< .001				
Studierende Engagement	S_1	0.474	< .001	4.7	0.99	0.69	0.76
	S_2	0.759	< .001				
	S_3	0.876	< .001				
Lernen Gesamtbewertung	L_1	0.857	< .001	4.48	1.13	0.94	0.94
	L_2	0.887	< .001				
	L_3	0.796	< .001				
	G_1	0.899	< .001				
	G_2	0.841	< .001				
	G_3	0.818	< .001				
CFA Fit-Indikatoren	CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		
	0.929	0.917	0.0476	0.0631	Untere	Obere	
					0.0613	0.065	

Um die explorativ gefundene und auf Reliabilität geprüfte Faktorenstruktur weiter zu validieren (faktorielle Validität), eignet sich die konfirmatorische Faktorenanalyse (CFA) (Gäde et al., 2020). Hierbei wird die angenommene Faktorenstruktur (wie in Tabelle 5 abgebildet) vorgegeben und getestet. Zur Beurteilung der modellierten Faktorenstruktur stehen Fit-Indikatoren zur Verfügung. Diese vergleichen das postulierte Modell (z.B. T_1, T_2, T_3 und T_4, bilden einen Faktor, während T_5, T_6, T_7 einen anderen Faktor bilden) mit der tatsächlichen korrelativen Struktur der Daten und geben an, wie gut das postulierte Modell und die Datenstruktur zusammenpassen. Dabei wird nicht nur (wie bei der

Reliabilitätsanalyse) die interne Konsistenz der einzelnen Skala überprüft, sondern auch ob sich die Faktoren voneinander abgrenzen (Diskriminante Validität). Es existiert eine Vielzahl an Fit-Indikatoren zur Überprüfung von Faktorenanalysen und Strukturgleichungsmodellen. Weiber und Mühlhaus (2014, S. 201) listen in ihrem Überblick beispielsweise 22 unterschiedliche Indikatoren auf. Große Einigkeit besteht darüber sich nicht auf einen Fit-Indikator zu verlassen, sondern das Modell mit mehreren Indikatoren zu testen, die aus möglichst unterschiedlichen „Familien“ kommen sollten. Häufig werden absolute Fit-Indikatoren und inkrementelle Fit-Indikatoren unterschieden (Berning, 2019, S. 8; Dash & Paul, 2021, S. 2; Weiber & Mühlhaus, 2014, S. 222). Absolute Fit-Indikatoren „beschreiben die Abweichung von einem perfekten Modell“, während inkrementelle Indikatoren „ein Kontinuum von null für das schlechteste Modell bis eins für ein perfektes Modell“ beschreiben (Berning, 2019, S. 8). Bei absoluten Fit-Indikatoren beschreiben Werte nahe null ein gutes Modell, da die Abweichung zum perfekten Modell möglichst klein sein sollte. Inkrementelle Indikatoren sollten Werte nahe an 1 aufweisen, da dies eine geringe Abweichung zum perfekten Modell beschreibt. Zur Interpretation der CFA werden, wie in Tabelle 6 dargestellt, vier Indikatoren (zwei inkrementelle und zwei absolute Indikatoren) verwendet und anhand der angegebenen Schwellenwerte interpretiert. Alle vier Werte weisen einheitlich auf eine solide Faktorenstruktur des Modells hin. Für TLI, CFI und SRMR unter- bzw. überschreiten die gemessenen Werte die Schwellenwerte jeweils. Für RMSEA liegt der Wert mit 0,0631 etwas über dem idealen Wert von $\leq 0,05$ aber im akzeptablen Bereich bis $< 0,08$. Auch das obere Ende des 90% Konfidenzintervalls liegt unter der Schwelle von 0,08, weshalb auch mit Blick auf den RMSEA von einer soliden faktoriellen Struktur ausgegangen werden kann.

Tabelle 6: Konfirmatorische Faktorenanalyse: Fit-Indikatoren und Schwellenwerte (N=3067)

	Fit-Indikator	Kurz	Wert	Schwellenwerte
Inkrementelle Fit-Indikatoren	Tucker Lewis Index	TLI	0,917	≥ 0,90
	Comparative Fit Index	CFI	0,929	≥ 0,90
Absolute Fit-Indikatoren	Standardized Root Mean Square Residual	SRMR	0,0476	≤ 0,10
	Root Mean Square Error of Approximation	RMSEA	0,0631	≤ 0,05-0,08
		RMSEA Konfidenzintervall	RMSEA 90% CI Untere 0,0613	Obere 0,065

Als Ergänzung zur faktoriellen Validität eines Konstrukts kann seine konvergente Validität überprüft werden. Damit wird überprüft, inwiefern Daten mit anderen Daten korrelieren, die das gleiche (oder zumindest ein ähnliches) Merkmal erheben. Ist dies der Fall, so kann davon ausgegangen werden, dass ein Konstrukt tatsächlich das misst, was es zu messen behauptet, also gültig ist.

Zur Überprüfung der konvergenten Validität wurden in 11 Veranstaltungen zusätzliche Daten mit dem MEEGA+ Fragebogen erhoben (Petri et al., 2017). Es handelt sich hierbei um eines der zuvor erwähnten Instrumente zur Evaluation von Lernspielen, mit dem sich teilweise Überschneidungen zu dem selbst entwickelten ZMS-Fragebogen ergeben. So erhebt MEEGA+ beispielsweise Items zur sozialen Interaktion während des Spiels, was über inhaltliche Schnittmengen mit Team Atmosphäre verfügen dürfte. Zudem fragt MEEGA+ ebenfalls die Zufriedenheit („Satisfaction“) ab. Das Lernen wird in MEEGA+ (ähnlich wie z.B. im Bielefelder Modell der Lernzielorientierten Evaluation) über ausformulierte lernzielspezifische Items erfasst. In Tabelle 7 sind inhaltlich ähnliche Skalen von MEEGA+ und dem ZMS-Fragebogen sowie deren Korrelation dargestellt.

Es zeigen sich schwache bis sehr hohe Korrelationen zwischen den Skalen der beiden Erhebungsinstrumente. Team Atmosphäre korreliert höher mit der Social Interaction Skala als Team Aufgabenverteilung. Dies ist ein inhaltlich nachvollziehbarer Unterschied, da die Items für die Social Interaction Skala eher atmosphärische Aspekte beleuchten (z.B. „I felt good interacting with other players during the game.“) während Fragen der Arbeitsorganisation gar nicht angesprochen werden. Die Korrelation zwischen Planspiel Verständnis und Learnability weist in die gleiche Richtung, fällt jedoch eher gering aus. Planspiel

Verständnis fragt mit zwei Items, ob das Spiel verstanden wurde, während Learnability eher den Prozess des Lernens abfragt (z.B. „I needed to learn a few things before I could play the game.“ oder “I think that most people would learn to play this game very quickly.”). Dass die beiden Skalen korrelieren ist inhaltlich nachvollziehbar und zu erwarten. Die eher geringe Korrelation ist ebenfalls nachvollziehbar, da die inhaltliche Überschneidung der beiden Skalen zwar vorhanden, aber keinesfalls deckungsgleich ist. Sehr hoch sind die Korrelationen zwischen Lernen und Gesamtbewertung und den MEEGA+ Skalen, die ebenfalls die allgemeine Zufriedenheit oder das Lernen abfragen. Hier ist die höchste Übereinstimmung zwischen den beiden Erhebungsinstrumenten zu erwarten, da sie an dieser Stelle explizit auf den gleichen Inhalt fokussieren: Zufriedenheit und Lernen.

Tabelle 7: Konvergente Validität, Vergleich mit MEEGA+

ZMS	MEEGA+	Pearsons r	P	N
Team Atmosphäre	Social Interaction	0.531	< .001	157
Team Aufgabenverteilung	Social Interaction	0.385	< .001	157
Planspiel Verständnis	Learnability (wie gut/leicht das Spiel erlernt werden)	0.231	0.004	157
Lernen und Gesamtbewertung	Satisfaction	0.774	< .001	157
Lernen und Gesamtbewertung	Learning Topic (spezifische Lernzielitems)	0.74	< .001	132
Lernen und Gesamtbewertung	Learning (allgemeine Formulierung)	0.764	< .001	165

Wie oben erwähnt werden im Rahmen der Evaluation am ZMS nicht nur Studierende befragt, ein kürzerer Fragebogen richtet sich auch an Lehrende. Dies bietet eine weitere Möglichkeit zur Analyse der konvergenten Validität, da teilweise überschneidende Sachverhalte bei Studierenden und Lehrenden abgefragt werden. Die Korrelationen zwischen den Antworten von Studierenden und Lehrenden wurden für Studierende mit Veranstaltungsmitteln gerechnet. N bezieht sich in Tabelle 8 also auf Veranstaltungen.

Tabelle 8: Konvergente Validität: Vergleich Lehrende und Studierende

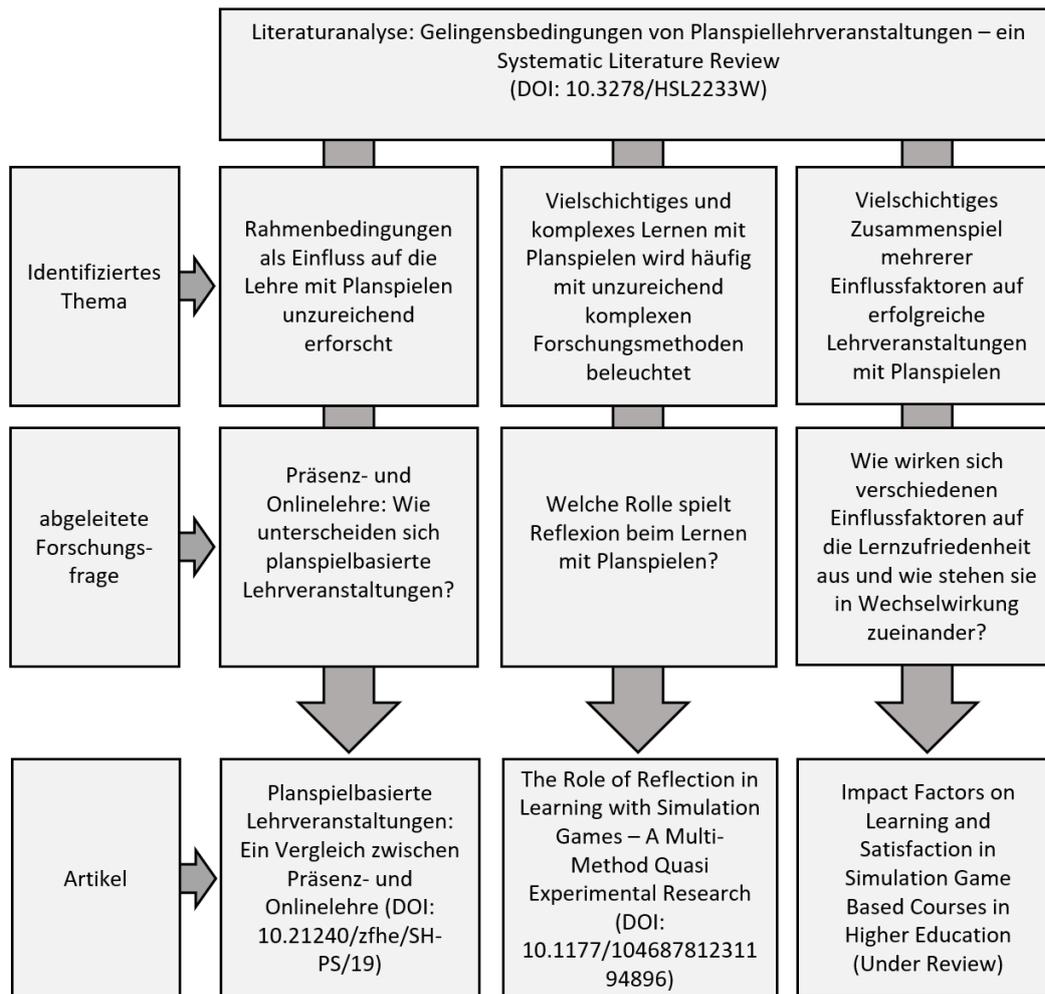
Studierende	Lehrende	Pearsons r	P	N
Konstrukt: Lernen und Gesamtbewertung	Item: Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?	0.381	<.001	115
Item: Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?	Item: Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?	0.366	<.001	115
Konstrukt: Studierende Engagement	Konstrukt: Studierende Engagement	0.171	0.033	116

Mittelstarke höchst signifikante Korrelationen (Zeilen 1 und 2, Tabelle 8) zeigen, dass Lehrende und Studierende hinsichtlich der Gesamtbeurteilung übereinstimmen. Ebenfalls signifikant und positiv, aber eher schwach ausgeprägt, ist die Korrelation zwischen Lehrenden und Studierenden im Hinblick auf das Engagement der Studierenden (Zeile 3, Tabelle 8). Tendenziell weisen die Urteile von Lehrenden und Studierenden in die gleiche Richtung, was für die Validität der studentischen Antworten spricht. Auch andere Untersuchungen mit dem ZMS-Fragebogen weisen in diese Richtung. Beispielsweise schätzten Studierende diejenigen Dozierenden als kompetenter im Umgang mit einem Planspiel ein, die das Planspiel in den letzten Jahren tatsächlich häufiger durchgeführt hatten (F. Trautwein & Alf, 2024, S. 54).

Der entwickelte und für diese Arbeit verwendete ZMS-Fragebogen, so lässt sich schlussfolgernd sagen, verfügt über eine ausreichende Reliabilität und Validität. Anhand der Daten zur internen Konsistenz (gemessen mit Cronbachs Alpha und McDonalds Omega) kann auf reliable Skalen geschlossen werden. Diese wurden anhand explorativer Faktorenanalysen entwickelt (F. Trautwein & Alf, 2023) und konfirmatorisch überprüft. Die ausgewählten Fit-Indikatoren weisen auf ein valides Modell hin. Der Vergleich mit einem Messinstrument, das ähnliche Aspekte abbildet, lässt auf konvergente Validität des Fragebogens schließen. Ein Vergleich zwischen der Einschätzung von Lehrenden und Studierenden kommt für die Gesamtbewertung und das Engagement der Studierenden zu einem übereinstimmenden Ergebnis.

Der hier dargestellte Fragebogen für Studierende kommt in allen empirischen Studien im Rahmen dieser Arbeit zum Einsatz.

Abbildung 5: Literaturstudie und daraus abgeleitete empirische Artikel



Als eine zentrale Forschungslücke wurden die Rahmenbedingungen als Einflussfaktor auf die Lehre mit Planspielen identifiziert: „Keine der ausgewerteten Studien nimmt den Einfluss von Rahmenbedingungen (z. B. Kursgröße, räumliche Ausstattung, Freiwilligkeit, Verbindung mit Prüfung) auf die Lehre mit Planspielen in den Blick.“ (Alf, 2022, S. 476). Dieser blinde Fleck liegt vermutlich in der üblichen Vorgehensweise der bisherigen Forschung begründet, in der zumeist untersucht wird, inwiefern ein bestimmtes Lernziel mit einem bestimmten Planspiel erreicht werden kann. Dieser fokussierte Blick eröffnet kaum Möglichkeiten den Einfluss von Rahmenbedingungen in das Forschungsdesign zu integrieren, da hierfür die quasiexperimentelle Einführung unterschiedlicher Bedingungen notwendig wäre (z.B. unterschiedliche Räumlichkeiten, unterschiedliche Kursgrößen etc.). Zur Erhebung von Rahmenbedingungen bot die Corona-Pandemie eine einmalige Gelegenheit. Planspielbasierte Lehrveranstaltungen konnten in Präsenz- und als Distanzveranstaltung miteinander verglichen werden. Diesem Thema widmet sich

der zweite Artikel, Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre (Alf & Trautwein, 2023).

Die Literaturanalyse weist zudem darauf hin, dass das vielschichtige und komplexe Lernen mit Planspielen häufig mit Forschungsmethoden beleuchtet wird, die dieser Komplexität nicht gerecht werden: „Wünschenswert für die zukünftige Forschung sind triangulative Ansätze, die der Komplexität des Lernprozesses anhand ihrer methodischen Vielfalt eher entsprechen.“ (Alf, 2022, S. 476). Mit dem Artikel *The Role of Reflection in Learning with Simulation Games – A Multi-Method Quasi Experimental Research* wird dieser Kritik an der bisherigen Forschung Rechnung getragen. Der Artikel untersucht, inwiefern Reflexionsangebote den Lernprozess fördern. Hierfür werden Spieldaten, qualitative Daten und quantitative Daten zu verschiedenen Zeitpunkten erhoben und für die Analyse in Beziehung zueinander gesetzt.

Sodann macht die Literaturstudie deutlich, dass das Gelingen planspielbasierter Lehre nicht von einzelnen Aspekten abhängt, sondern vielmehr von ihrem Zusammenspiel: „...dies weist auf den multifaktoriellen Charakter der Lehr-Lernforschung (mit Planspielen) hin. Es ist das maßvolle Zusammenspiel der Faktoren, das zum Gelingen planspielbasierter Lehrveranstaltungen beiträgt. Isoliert determinierende Faktoren zu benennen, würde der Komplexität der Methode nicht gerecht.“ Dieses Fazit aus der Literaturstudie wird im vierten Artikel adressiert. In einem Strukturgleichungsmodell wird das Zusammenspiel einer Vielzahl von Einflussfaktoren empirisch dargestellt und modelliert. Auch das Zusammenwirken von Einflussfaktoren (indirekte Effekte) auf den Lehrerfolg kann so modelliert werden.

Nach dieser Einführung folgen in den nächsten Unterkapiteln die jeweiligen Artikel, wie sie veröffentlicht oder in einem Fall zur Veröffentlichung eingereicht sind. Um ein stimmiges Gesamtmanuskript zu erhalten wurde lediglich das Layout der Artikel vereinheitlicht (z.B. Schriftart, Schriftgröße, leichte optische Veränderungen an Tabellen). Die Literaturverzeichnisse am Ende der Artikel beziehen sich jeweils ausschließlich auf den entsprechenden Artikel. In den Artikeln zitierte Quellen stehen nicht noch einmal im Literaturverzeichnis ganz am Ende dieses Rahmentextes. Taucht eine Quelle aus einem Artikel doch im Literaturverzeichnis ganz am Ende des Rahmentextes auf, so nur deshalb, weil die Quelle auch im Rahmentext verwendet wird. Drei der vier Artikel wurden zusammen mit anderen

Autorinnen und Autoren verfasst. Die leicht ausgegrauten Textpassagen weisen darauf hin, dass diese nicht von mir selbst verfasst wurden.

5.1 Gelingensbedingungen von Planspiellehrveranstaltungen – Ein Systematic Literature Review

Abstract

Planspiele werden in der Hochschullehre vielfach eingesetzt. Welche Merkmale aber tragen zu einer gewinnbringenden Lehre mit Planspielen bei? Anhand einer systematischen Literaturstudie wird diese Frage untersucht. Der Überblick umfasst 58 empirische Forschungsarbeiten und diskutiert deren Erkenntnisse unter den Gesichtspunkten Qualität des Planspiels, Einfluss der Lehrenden, Einfluss der Studierenden sowie im Hinblick auf den Lehrerfolg. Es zeigt sich, dass Planspiele eine gewinnbringende Lernmethode sind, mit der sehr unterschiedliche Lernfelder adressiert werden. Das Planspiel sollte im Hinblick auf Daten und Wirkungsverläufe realistisch sowie an die Bedürfnisse von Lernenden anpassbar sein. Lehrende nehmen eine zentrale Rolle ein. Erst durch Einführung, Anpassung, Erklärung, Feedback und schlussendlich Auswertung der Simulation, durch die Spielleitung, entsteht für Lernende ein Mehrwert. Hinsichtlich der Studierenden wurden insbesondere das persönliche Engagement und die Teamkommunikation als zentrale Einflussfaktoren identifiziert.

DOI: [10.3278/HSL2233W](https://doi.org/10.3278/HSL2233W)

Autor: Tobias Alf

Einleitung

An vielen Hochschulen sind Planspiele fester Bestandteil der Lehre. Zudem werden sie vermehrt auch in schulischen und betrieblichen Lernkontexten eingesetzt. Planspiele können als Spielumgebung beschrieben werden, in der Aspekte der Wirklichkeit modellhaft nachgebildet werden (Simulation). In dieser Umgebung übernehmen Lernende Rollen und agieren, unter Berücksichtigung von Spielregeln, als diese, um simulierte Probleme zu bearbeiten (Kriz, 2011). Ebenfalls kennzeichnend für Planspiele ist der Wechsel von Briefing (Input, Spieleinführung), Spiel (Entscheidungen treffen, Handeln) und Debriefing (Reflexion der Erfahrungen und Übertragung auf Wirklichkeit) (Schwägele et al., 2021). Häufig werden Planspiele mit dem Ziel eingesetzt, Komplexität und systemische Effekte begreifbar zu machen (Imhof & Starker, 2020; Kriz, 2009; Pasin & Giroux, 2011). In der Literatur werden Planspiele als problemorientierte (Mandl, 2004) und erfahrungsorientierte Lehr-Lernmethode diskutiert (Geithner & Menzel, 2016; Schwägele, 2015). Diese Einordnung entspricht dem charakteristischen Wechsel von Aktivität/Erfahrung einerseits und Reflexion der Erfahrung andererseits. Obwohl es relativ viele Publikationen rund um das Thema Lernen mit Planspielen gibt, ist wenig über die strukturellen Gelingensbedingungen der Lehre mit Planspielen bekannt. Welche Aspekte müssen wie zusammenspielen, damit die Lehre mit Planspielen gut gelingt? Mit der Methode Systematic Review werden hier Forschungsarbeiten zum Lernen mit Planspielen hinsichtlich der Frage ausgewertet, unter welchen Voraussetzungen die Lehre mit Planspielen gelingen kann. Als grobe inhaltliche Strukturierung wird das multifaktorielle Modell der Lehrveranstaltungsevaluation nach Rindermann herangezogen (2009), wonach sich der Lehrerfolg durch das Zusammenspiel von Rahmenbedingungen, Dozierenden und Studierenden ergibt. In der planspielbasierten Lehre wird der Art des eingesetzten Planspiels eine große Bedeutung für den Lehr-Lernprozess zugesprochen. Daher wird das Modell um den zusätzlichen Einflussfaktor Planspiel ergänzt. Die konkreten Fragen lauten:

- Welchen Einfluss haben die Rahmenbedingungen auf den Lehrerfolg?
- Welchen Einfluss hat das Planspiel auf den Lehrerfolg?
- Welchen Einfluss hat die Planspielleitung (Dozierende) auf den Lehrerfolg?
- Welchen Einfluss haben Studierende auf den Lehrerfolg?
- Lehrerfolg: Wie gut gelingt Lernen mit Planspielen?

Zur Methode dieser Arbeit

Ein Systematic Review ist eine wissenschaftliche Literaturarbeit, mit dem Ziel, einen Überblick über die Forschung in einem bestimmten Feld zu geben. Im Voraus wird festgelegt, welche Literatur in Frage kommt und wo sie nach welchen Kriterien gesucht wird. Es handelt sich um eine systematische Vorgehensweise der Literatursuche, im Gegensatz zu einer eher zufälligen Vorgehensweise (Gough et al., 2012, S. 2). Je nach Quelle werden für Systematic Reviews von der Entwicklung der Forschungsfrage bis zur Veröffentlichung der Ergebnisse verschieden detaillierte Schritte vorgeschlagen (Gough et al., 2012; Gough et al., 2017). Das methodische Vorgehen dieser Arbeit entspricht folgenden Schritten, die sich an den Vorschlägen von Gough et al. orientieren:

- Entwicklung der Forschungsfrage, wie oben dargestellt
- Iterative Entwicklung einer Suchstrategie incl. Suchstrings
- Auswahl relevanter Literaturdatenbanken für die Suche
- Selektionskriterien: Entwicklung von Einschluss- und Ausschlusskriterien
- Sichtung von Titel und Abstract und Auswahl entsprechender Publikationen anhand von Selektionskriterien
- Lektüre, Kodierung und Bewertung der eingeschlossenen Studien
- Zusammenfassen von Ergebnissen anhand Kodierung

Die Entwicklung der Suchstrings erfolgte in einem iterativen Prozess, um angemessen breite, aber nicht ausufernde Ergebnisse zu erhalten. So erzielt z.B. die Suche nach Game oder Spiel ohne Einschränkung durch Simulation oder Plan- eine kaum zu überblickende Anzahl an Artikeln, die sehr heterogene Spiele thematisieren. Schlussendlich wurde am 20.10.2020 mit folgenden Suchstrings gearbeitet:

Fachportal Pädagogik: (Planspiel OR Simulationsspiel OR „Simulation Game“ OR “Simulation Games” OR Simulation-Game OR Simulation-Games OR Simulationgame OR Simulationgames) AND Qualität* OR Quality* OR Effekt* OR Effect* OR Evalu* OR Auswirkung* OR Lehrevaluation OR Lernzielerreichung OR Lernen OR Achievement OR “Learning Goal” OR “Learning Target”)

ERIC: ("simulation game" OR "simulationgame" OR "simulation-game") AND (quality OR evaluation OR effect OR result OR learning OR learning target OR achievement OR learning goal)

Auch die Auswahl der Literaturdatenbanken erfolgte iterativ. Um die große Anzahl an Veröffentlichungen bearbeiten zu können, beschränkt sich die Suche auf die deutsch- und englischsprachige Datenbank mit den relevantesten Treffern, nämlich das Fachportal Pädagogik und das Educational Resources Information Center (ERIC). Als Suchstrategie wurde in beiden Datenbanken eine Freitextsuche mit Booleschen Operatoren verwendet, um die Themenkomplexe, Planspiel einerseits und Qualitätsmerkmale/Gelingensbedingungen andererseits, einzugrenzen. Um die Suchergebnisse weiter einzugrenzen wurden weitere Selektionskriterien festgelegt: Es kommen für diese Arbeit nur Beiträge aus Zeitschriften mit Reviewverfahren aus den letzten 10 Jahren in Betracht. Zudem müssen die Artikel eine empirische Herangehensweise beinhalten (=qualitative oder quantitative Datenerhebung und Analyse). Nicht in Frage kommen Systematic Reviews, theoretische Arbeiten und Artikel die Planspiele im Sinne eines Best Practice vorstellen. Zudem müssen die Artikel pädagogisch/didaktische Fragestellungen der Lehre mit Planspielen adressieren, im Gegensatz zu Fragen des Designs oder der technischen Programmierung von Planspielen.

Die Suche in den genannten Datenbanken mit den angegebenen Suchstrings ergibt nach Abzug von Dubletten 257 Titel. Nach Sichtung von Titel und Abstracts (und Anwendung der beschriebenen Selektionskriterien) kommen 70 Artikel für die weitere Analyse in Betracht (Tabelle 1).

Tabelle 1: Auswahl relevanter Artikel

Fachportal Pädagogik	143
ERIC	+181
Gesamt	=324
Dubletten (Artikel in beiden Datenbanken)	-67
Zu sichtende Artikel	=257
Relevant nach Sichtung von Titel und Abstract	
Ausschlussgründe:	
- Fälschlicherweise als Peer-Review in Datensatz	70
- „Simulation“ als unklarer Begriff	
- Kein Planspiel, sondern Videospiele oder Single-Player-Spiel	
- Fokus auf technische Fragestellung/Programmierung/Grafik	
Relevant nach Lektüre	
Ausschlussgründe:	
- Fälschlicherweise als Peer-Review in Datensatz	58
- „Simulation“ als unklarer Begriff, kein Planspiel	
- Eher technische Fragestellung	

Für die inhaltliche Auswertung (Lektüre, Kodierung und Bewertung) der relevanten Veröffentlichungen wurden entsprechende Passagen in den Texten in einem zweistufigen Verfahren kategorisiert. Zunächst wurden die Passagen den oben genannten Faktoren (Rahmenbedingungen, Studierende, Lehrende, Planspiel und Lernerfolg) zugeordnet. In einem zweiten Schritt wurden innerhalb der Faktoren induktiv Kategorien gebildet, um einen besseren Überblick über die Ergebnisse zu erhalten. Die Ergebnisse zum Lernerfolg wurden nach Studiendesign kategorisiert, um eine bessere Vergleichbarkeit der Ergebnisse zu ermöglichen. Die Ergebnisse aus den verschiedenen Forschungen werden in dieser Arbeit neu organisiert und zueinander in Beziehung gesetzt. Das methodische Vorgehen entspricht daher eher einem konfigurativen als einem aggregativen Ansatz (vgl. Gough et al., 2017). Während des Prozesses der Lektüre und Kategorisierung wurde die Literatur zudem einer kritischen Bewertung unterzogen, insbesondere im Hinblick auf die Fragestellung dieser Arbeit. Hierbei wurden zwölf weitere Veröffentlichungen ausgeschlossen (Tabelle 1), zumeist weil sie keine Planspiele (im oben beschriebenen Sinne) behandeln, oder weil technische Fragestellungen im Mittelpunkt standen. Schlussendlich finden Ergebnisse aus 58 Arbeiten Eingang in die weitere Analyse.

Wie gelingt gute Lehre mit Planspielen?

In diesem Kapitel werden Gelingensbedingungen der Lehre mit Planspielen aus drei Perspektiven diskutiert: Was trägt das Planspiel an sich zum Gelingen der Lehre bei? Wie kann die Spielleitung (Dozierende) zum Erfolg der Lehre beitragen und welchen Beitrag leisten Lernende? Der vierte Aspekt, die Rahmenbedingungen (z.B. räumliche Ausstattung, Anzahl teilnehmende Studierende, Zeitpunkt und Zeitraum etc.), wurde in keinem der untersuchten Artikel untersucht und fällt daher auch hier weg. Sodann wird dargestellt, welche Lernerfolge in den relevanten Artikeln präsentiert und diskutiert werden.

Planspiel

Werden Planspiele ins Zentrum des Lehr-Lernprozesses gestellt, so dienen sie in zweifacher Hinsicht als Lernmethode. Studierende sammeln Erfahrungswissen, indem sie als Akteur in einem bestimmten Feld (z.B. Geschäftsführung, politisches

Amt) handeln und mit den Konsequenzen ihres Handelns umgehen müssen. Sodann dient das Planspiel als Anker für anschließende Reflexionen. Das Entwickeln von Planspielen ist ein hoch komplexer Prozess. Lerninhalte und Simulation der Realität müssen iterativ in Spieldesign und Spielmechanismen überführt werden (Duke, 1980; Herzig, 2019). Wie aber muss ein Planspiel beschaffen sein, um Lernprozesse zu fördern? 13 der untersuchten Studien liefern explizite Erkenntnisse zur Beantwortung dieser Frage und werden im Folgenden dargestellt. Die Ergebnisse lassen sich in vier Kategorien einteilen: Spielgestaltung, Anpassbarkeit, Anleitung und Feedback sowie Authentizität.

Unter dem Gesichtspunkt Spielgestaltung wird einheitlich betont, dass Planspiele für Lernende leicht zu bedienen und nachvollziehbar in der Handhabung sein sollten (Abdullah et al., 2013, 105; Dick & Akbulut, 2020, 630; Matute-Vallejo & Melero-Polo, 2019, 80). Dem zu widersprechen scheint die Aussage, dass Planspiele vielfältig und unvorhersehbar sein sollten (Lin & Tu, 2012, 1167), also gerade nicht leicht und nachvollziehbar. Es zeigt sich jedoch, dass erstere im Wesentlichen die Nutzungsfreundlichkeit (Spielmaterialien, Bedienbarkeit, Regeln, technische Handhabung etc.) von Planspielen in den Blick nehmen und zu dem Schluss kommen, dass diese einfach, logisch und nachvollziehbar sein muss. Lin und Tu hingegen geht es um Inhalte und Szenarien. Sie betonen, dass diese komplex, unvorhersehbar und vielfältig sein sollen. Ein gutes Planspiel zeichnet sich folglich dadurch aus, dass es intuitiv und leicht zu bedienen ist, gleichzeitig aber komplexe, unvorhersehbare und vielfältige Inhalte und Erfahrungen bereit hält.

Die Anpassbarkeit von Planspielen meint die situative Adaption eines Planspiels an die Gruppe der Lernenden. Dies wird sowohl im Hinblick auf Kultur/Sprache (Kwon et al., 2013, 249f.) als auch im Hinblick auf das Vorwissen der Lernenden (Brazhkin & Zimmerman, 2019, 378) betont. Da Planspiele in bestimmten Kulturen und Sprachen (häufig im angloamerikanischen Raum) designt werden, transportieren sie immanente kulturelle Werte und deren Sprache. Für Lernende aus anderen Kulturen müssen sie ggf. erklärt und angepasst werden. Ähnliches gilt für das Vorwissen der Lernenden. Je nach individuellem Spielerfolg könnten Planspiele automatisch komplexer (weniger Information, zusätzliche Aufgaben) oder einfacher (mehr Hilfestellung, reduzierte Schwierigkeit) werden. Anleitung und Feedback meint, dass Lernende im Spielverlauf durch das Spiel selbst Rückmeldung über ihren Spielfortschritt und Spielerfolg erhalten. Dies führt zu

einem höheren Engagement und einer intensiveren Auseinandersetzung mit den Spielinhalten (Buil et al., 2018, 189f.; Sandberg et al., 2012, 848). Da Planspiele einen Ausschnitt der Realität vereinfacht darstellen sollen, kommt der Authentizität von Planspielen eine wichtige Rolle zu. Ney et al. (2014, 136) und Misfeldt (2015, 187) kommen übereinstimmend zu dem Schluss, dass Authentizität nicht durch eine möglichst echte sinnliche Inszenierung entsteht, sondern durch realistisch simulierte Konsequenzen und Wirkungsverläufe.

Spielleitung

Bei der Durchführung von Planspielen hat die Spielleitung mehrere zentrale Funktionen inne. Sie ist ggf. bei der Auswahl der Simulation beteiligt, hat die Aufgabe ins Planspiel einzuführen (Briefing), den Spielprozess zu steuern und ist im Debriefing dafür zuständig, den Austausch über das Lernen zu moderieren. Aus zwölf Studien werden im Folgenden Erkenntnisse diskutiert, die dem Einflussfaktor Spielleitung zugeordnet werden können.

Die Ergebnisse aus dem Bereich Spielleitung lassen sich in drei Kategorien gliedern: Person, Feedback und Struktur sowie Debriefing.

Zwei Studien betonen, relativ allgemein, die Wichtigkeit einer kompetenten Spielleitung (Kategorie Person) für den Lernprozess mit Planspielen (Dick & Akbulut, 2020, 630; Mayer et al., 2013, 97). In einem Strukturgleichungsmodell zeigen Mayer et al. (ebd.), dass die Spielleitung mehr Einfluss auf die Lernzufriedenheit nimmt als das Planspiel selbst. Eine weitere zentrale Anforderung an die Person der Spielleitung ist die Übernahme von Rollen (deNoyelles & Raider-Roth, 2016, 346). Anders als in klassischen Lehrformaten tritt die Spielleitung (bei vielen Planspielen) nicht nur als Lehrperson auf, sondern auch als Spielcharakter (Bank, politische Institution etc.), was ein schnelles und glaubhaftes Springen zwischen den Rollen erfordert. Warum die Spielleitung in der Lehre mit Planspielen eine so zentrale Bedeutung hat, erklärt sich durch die Kategorie Feedback und Struktur. Dies meint die Begleitung der Studierenden während des Spielprozesses in einer solchen Weise zu gestalten, dass Lernende einerseits herausgefordert sind und sich ausprobieren müssen, andererseits aber so viel Feedback und Struktur erhalten, dass sie das Spiel nicht als willkürlich oder überfordernd erleben (Abdullah et al., 2013, 103; Buil et al., 2018, S. 189). In der Planspielliteratur wird diese Begleitung häufig als Facilitation beschrieben (Leigh & Spindler, 2005). Das Abwägen zwischen mehr oder weniger Feedback und

Struktur erfordert ein hohes Maß an Wahrnehmung und Sensibilität für die jeweilige Situation. Als Spielleitung moderieren Lehrende zwischen der Komplexität und Ungewissheit des Planspiels einerseits und den Bedürfnissen der Lernenden andererseits. Erst durch das richtige Maß an Feedback und Struktur während des Spielverlaufs wird ein Planspiel für Studierende zu einer Lernerfahrung mit Mehrwert. Eine weitere Kategorie, die der Spielleitung zugeordnet werden kann, ist das Debriefing, also die angeleitete Reflexion des im Planspiel Erlebten und die Übertragung auf die Realität. In vier Studien werden Forschungsergebnisse zum Debriefing berichtet, wobei zwei dieser Studien (Beranic & Hericko, 2019; Richards & Camuso, 2015) hierbei auch die Rolle der Studierenden hervorheben: Debriefings sollten so organisiert werden, dass Lernende von ihren Erkenntnissen berichten und wechselseitig voneinander profitieren können. Auf diese Weise gewinnt die Spielerfahrung an Reichhaltigkeit und Ganzheitlichkeit. Zudem können auch Lerninhalte besprochen werden, die Studierende aus dem Spiel ziehen, die aber ggf. von Lehrenden gar nicht intendiert waren. Dies wirft die Frage auf, inwiefern Lehrende den in vielen Planspielen ohnehin bestehenden Wettbewerbscharakter verstärken sollten. Ist dieser stark ausgeprägt, werden Erkenntnisse und Erfahrungen womöglich aus spielstrategischen Gründen nicht offenbart – zumindest nicht in den Debriefings zwischen den Spielrunden.

Während das Lehrendenverhalten in der Kategorie Struktur und Feedback sehr situativ und individuell ist, lässt sich das Debriefing mit umfangreichen Reflexions- und Moderationsmethoden vorbereiten und gezielt gestalten. Es wäre daher aufschlussreich unterschiedliche Methoden empirisch zu vergleichen. Keine der untersuchten Arbeiten, geht dieser Frage jedoch strukturiert nach. Dies ist eine erhebliche Lücke in der Lehr-Lernforschung zu Planspielen, obwohl das Debriefing zur zentralen Aufgabe der Planspielleitung zählt.

Studierende

Die Forschung zu planspielbasierten Lehr-/Lernveranstaltungen scheint sich auf Studierende zu konzentrieren. Diesem Einflussfaktor lassen sich Ergebnisse aus 33 Studien zuordnen. Diese lassen sich in die Kategorien Teamarbeit, Engagement, Emotion und Erleben sowie Persönlichkeitsmerkmale und Biografie gliedern.

Acht Studien beschäftigen sich mit der Rolle der Teamarbeit im Rahmen von Planspielen. Hierbei zeigen sich sowohl die Schwierigkeiten als auch die Chancen von Teamarbeit: Zusätzlich zu den Inhalten des Planspiels wird die Teamarbeit (Organisation, Absprachen etc.) als eigene Aufgabe beschrieben (Brazhkin & Zimmerman, 2019, 369), die Konflikte und zusätzliche Schwierigkeiten mit sich bringt (Scholkmann et al., 2017, 664f.; Siewiorek et al., 2012). Die Mehrheit der Studien zeigt einen Zusammenhang zwischen guter Teamarbeit und Erfolg im Planspiel (M. I. Hwang, 2018; Siewiorek et al., 2013; Tao et al., 2012; Treen et al., 2016). Hierbei kommt es vor allem auf die qualitative Ausgestaltung der Teamarbeit an: Es sind intensive, kritische Gruppendiskussionen und konstruktive Konflikte (Nachfragen, Ergänzen, Widersprechen), die zu einer verbesserten Leistung führen (Tao et al., 2012, 1359; van den Bossche et al., 2011, 294). In der Kategorie Engagement sind Studienergebnisse zusammengefasst, die den persönlichen Einsatz der Studierenden in den Lernprozess beschreiben. Übereinstimmend berichten diese Studien, dass ein hohes Engagement mit einem höheren Lern- oder Spielerfolg einhergeht (Alas et al., 2018, 316; Buil et al., 2020, 306; Corbeil & Laveault, 2011, 469). In der Kategorie Emotion und Erleben finden sich Ergebnisse, die das persönliche Empfinden während des Spiel- und Lernprozesses beschreiben. Die Studien beschreiben überwiegend positive Emotionen im Verlauf von Planspielen (Abdullah et al., 2013, 105; Buil et al., 2018, 189; Matute-Vallejo & Melero-Polo, 2019, 80f.), aber auch Überforderung und Frustration aufgrund der Unsicherheit und Komplexität von Planspielen (Brazhkin & Zimmerman, 2019, 369; Herzig, 2019, 117f.). Unter Persönlichkeitsmerkmale und Biografie werden Studienergebnisse diskutiert, die sich mit spezifischen Merkmalen von Studierenden auseinandersetzen, die diese mit ins Planspiel einbringen (z.B. Vorbildung, Studienmotivation etc.). Drei Studien berichten über Zusammenhänge zwischen Studienmotivation bzw. Interesse am Thema einerseits und einer positiven Bewertung des Planspiels andererseits (Mayer et al., 2013, 96; Nguyen, 2015, 13; Oberle & Leunig, 2016, 233). Sowohl Oberle und Leunig als auch Mayer et al. weisen jedoch darauf hin, dass dieser Zusammenhang als gering einzuschätzen ist. Tao et al. finden sogar einen gegensätzlichen Zusammenhang. Sehr motivierte Studierende bewerten das Planspiel tendenziell schlechter. Sie vermuten, dass zielstrebige und lernfokussierte Studierende vom spielerischen Ansatz irritiert sein könnten und ihn als unnötigen Umweg betrachten (Tao et al., 2012, 1357). Insgesamt scheint es keinen eindeutigen Zusammenhang zwischen genereller Studienmotivation und

Motivation zur Teilnahme an einem Planspiel zu geben. Vier Studien besprechen Zusammenhänge zwischen Lernpräferenzen bzw. Lerntypen und der Zufriedenheit mit der Lernmethode Planspiel (Auman, 2011, 470; Mayer et al., 2013, 96; Piercy et al., 2012, 10; Tao et al., 2012, 1357). Hier zeigt sich ein uneinheitliches Bild. Während Auman und Tao et al. Zusammenhänge zwischen Lerntypen und dem Lernen mit Planspielen berichten, finden Piercy et al. und Mayer et al. keine Zusammenhänge. Bei der Interpretation ist Vorsicht geboten, da in allen Erhebungen Lerntypen untersucht werden, diese jedoch teilweise auf unterschiedlichen, theoretischen Modellen basieren. Auch wenn der persönliche Lerntyp von Studierenden einen Einfluss auf die Zufriedenheit mit der Lernmethode Planspiel hat, scheint er generell nicht so deutlich zu sein, dass er sich in allen Studien zeigt. Zudem wird die Einteilung von Personen in Lerntypen in der Literatur kritisch diskutiert (Pashler et al., 2008; Schäfer, 2017, S. 7). Die Aussagekraft dieser Ergebnisse bleibt demnach fraglich. Mehrere Studien untersuchen den Einfluss der Vorbildung von Lernenden auf die Teilnahme an Planspielen. In vier Studien kann gezeigt werden, dass Notendurchschnitte oder unterschiedliche Bildungsabschlüsse keinen Effekt auf das Lernen mit Planspielen haben (Alas et al., 2018, 318; Brennan & Vos, 2013, 265; Monk & Lycett, 2016, 752; Scholkmann et al., 2017, 665). In keiner der hier angesprochenen Studien stand der Einfluss der Vorbildung auf das Lernen mit Planspielen als Forschungsfrage im Zentrum, weshalb die Erkenntnisse mit einer gewissen Vorsicht interpretiert werden sollten. Dennoch ergibt sich ein einheitliches Bild. Planspiele können als Lernmethode angesehen werden, mit der Lernende unterschiedlicher Vorbildung gleichermaßen angesprochen und gefördert werden können.

Insgesamt zeichnen die referierten Studien ein einheitliches Bild. Studierende tragen mit ihrem Engagement während einer Planspielveranstaltung erheblich zum Gelingen dieser bei. Dies gilt sowohl für die Zufriedenheit mit der Lernsituation als auch für den Erfolg im Planspiel sowie für das erreichte Lernergebnis. Dieses einheitliche Ergebnis ist keineswegs überraschend. In Planspielen erhalten Studierende die Möglichkeit, in simulierten Umgebungen Erfahrungen zu sammeln und diese zu reflektieren. Je intensiver sie von dieser Möglichkeit Gebrauch machen, desto eher treten die gewünschten Ergebnisse ein.

Lehrerfolg oder wie gut gelingt Lernen mit Planspielen

Von den 58 untersuchten Artikeln weisen 45 Forschungsergebnisse hinsichtlich des Lernerfolgs bzw. des Lernens mit Planspielen aus. Diese Ergebnisse werden im Folgenden, geordnet nach ihrem Forschungsdesign, vorgestellt.

Aus vielen Studien können Erkenntnisse berichtet werden, die auf einfache quantitative Befragungen (in der Regel am Ende der Veranstaltung) zurückgehen. Im Fokus steht hierbei die Zufriedenheit mit der Lernmethode oder die Einschätzung der Studierenden über ihren Lernerfolg. Es zeigt sich, dass Studierende mit Planspielen sehr zufrieden sind (Dick & Akbulut, 2020, 626; Geithner & Menzel, 2016, 247; Oberle & Leunig, 2016, S. 233) und die Lernmethode als sehr motivierend empfinden (Herzig, 2019, 180; Hilliard, 2014, S. 119; Piercy et al., 2012, 110). Es wird deutlich, dass Studierende das Lernen mit Planspielen als hilfreich erleben, um ein tiefes, anwendungsorientiertes Verständnis der Lerngegenstände zu erwerben. Grundlage hierfür sind eine als positiv erlebte, hohe Eigenaktivität und ein positives emotionales Erleben. Zu erwähnen ist auch die inhaltliche Breite der untersuchten Planspiele. Sie reichen von Betriebswirtschaft/Entrepreneurship (Buil et al., 2020; Mohsen et al., 2019; Zulfiqar et al., 2019) über Projektmanagement (Geithner & Menzel, 2016), Politik (Dube, 2015; Eränpalo, 2014) bis hin zum Erlernen von Computersystemen (Beranic & Hericko, 2019; Chen et al., 2015).

Es finden sich acht Studien, die Ergebnisse zum Lernen mit Planspielen aus Pre/Post-Befragungen berichten. Auch diese Studien thematisieren Planspiele mit sehr unterschiedlichen Lerninhalten (EU-Politik, soziale Ungleichheit, BWL/Management, und ERP-Systeme). Sieben der acht Studien, in denen Erwartungen (pre) mit den subjektiven Lernerfolgen verglichen werden (post), zeigen eine positive Einschätzung (Geithner & Menzel, 2016; Hamzeh et al., 2017; M. Hwang, 2019; Norris, 2013; Oberle & Leunig, 2016; Seethamraju, 2011; Williams, 2015). Lediglich bei Mayer et al. (2013, 97) fällt die Vorbefragung positiver aus als die Nachbefragung, das Lernen bleibt hier also hinter den Erwartungen der Studierenden zurück. Mayer et al. diskutieren als Grund, dass die Spielumgebung weniger effektiv und ansprechend ist als erwartet, was sich auch mit der Wahrnehmung der Lehrpersonen decke (ebd.).

Im Folgenden werden Studien diskutiert, die den Lerneffekt von Planspielen objektiv messen. Hierunter fallen Tests, die Auswertung von Daten aus den Planspielen, aber auch das Festhalten der objektiv genutzten Lernzeit durch Videografie. Zunächst werden fünf Studien diskutiert, die den Lernerfolg einmalig erheben. Cronan et al. (2012, 474) verwendet zur Ermittlung des Lernerfolges durch das Planspiel eine Wissensabfrage und kommt zum Ergebnis, dass sich das Planspiel zum Erlernen von ERP-Systemen eignet. Drei Studien nutzen die Spielergebnisse als Maß des Lernerfolgs⁷ und kommen zu positiven Einschätzungen (Alas et al., 2018; M. I. Hwang, 2018; van den Bossche et al., 2011). Einen anderen Ansatz verfolgen Scholkmann et al. (2017), die mit videografischer Auswertung die Lernzeitnutzung beobachten und bei effektiv genutzter Lernzeit auf den Lernerfolg schließen. Obwohl Planspiele Lernenden viel Freiraum lassen, verbringen sie ca. 80% der zur Verfügung stehenden Zeit mit beobachtbaren Lernaktivitäten, was eher den Werten von eng strukturierten Unterrichtsformen entspreche (ebd., 666). Trotz ihrer unterschiedlichen methodischen Herangehensweisen an den Forschungsgegenstand, schätzen die fünf Studien den Lernerfolg mit Planspielen anhand objektiver Kriterien als positiv ein.

Aus sieben Studien können Erkenntnisse berichtet werden, die den Lernerfolg zu zwei Zeitpunkten (pre/post) mit objektiven Daten erhoben und (in der Regel) durch gepaarte T-Tests analysiert haben. Auch in diesen Studien finden sich Planspiele, die sehr unterschiedliche Lerngegenstände adressieren. Alle sechs Studien berichten in den Post-Erhebungen von besseren Ergebnissen als in den Pre-Erhebungen. Pasin und Giroux (2011, 1250f.) untersuchen den Lerneffekt anhand des Vermeidens typischer Fehler. Sie gehen von einem Lerneffekt aus, wenn Studierende in einer späteren Spielphase einen typischen Fehler nicht mehr begehen. Tatsächlich werden zwei der drei untersuchten typischen Fehler in der zweiten Spielhälfte vermieden (kleine bis mittlere Effektstärke). Der Effekt wird stärker, wenn man diejenigen Planspielgruppen herausrechnet, die typische Fehler bereits in der ersten Hälfte des Planspiels vermieden haben, ihr Wissen also schon aus anderen Kontexten mitbrachten. Auch in den fünf Studien, die den

⁷ Viele Studien messen den Lehrerfolg anhand von Spieldaten. Dieses Vorgehen ist jedoch nicht unumstritten. Vgl. hierzu Pasin und Giroux 2011, S. 1248; Brazhkin und Zimmerman 2019, S. 371f..

Lernerfolg anhand von Aufgaben oder Wissenstests erheben, zeigt sich ein positives Bild: Palmunen et al. (2013) finden einen deutlichen Lernzuwachs zwischen Pre- und Posterhebung. Die Studierenden konnten nach dem Planspiel 73% mehr betriebswirtschaftlich relevante Verbindungen herstellen. Bioglio et al. (2019, 463f.) finden nach dem Planspiel ein signifikant sensitiveres Datenschutzverhalten, wobei durch das Studiendesign nicht klar gesagt werden kann, ob dies ein Resultat des Planspiels oder des begleitenden Unterrichts ist. Die drei weiteren Studien, die den Lernerfolg mit Tests oder Wissensabfragen erheben (Brennan & Vos, 2013; Oberle & Leunig, 2016; Sandberg et al., 2012), finden mittlere bis sehr starke Effekte zwischen Pre- und Posterhebung.

In Studiendesigns, die einer Planspielgruppe (Test) eine Kontrollgruppe gegenüberstellen, kommen vier von fünf Studien zu einem positiven Ergebnis für die Testgruppe (Bachen et al., 2012; Bello et al., 2016; Nussbaum et al., 2015; Prihadi et al., 2018). Lediglich Mohd Radzi et al. (2019, 172f.) finden in ihrer Studie keinen Lerneffekt. Das untersuchte Planspiel ist hier als Ergänzung zu einem Logistikseminar am Ende des Semesters zu verstehen. Die Studierenden haben also während des Semesters in klassischen Lernformaten Wissen erworben und können dieses am Ende des Semesters in einem Planspiel vertiefen. Die Planspielgruppe konnte im Gegensatz zu den Studierenden, die „nur“ den Kurs besuchten, ihr Wissen nicht signifikant erweitern. Evtl. eignen sich Planspiele eher als Einstieg in ein Thema. Dafür finden sich Hinweise in der oben besprochenen Studie von Brennan und Vos, bei denen Studierende mit zunächst geringem Wissensstand sehr starke Fortschritte machten. Prihadi et al. untersuchen den Einfluss eines Planspiels auf die Resilienz der Teilnehmenden und finden im Gegensatz zur Vergleichsgruppe einen positiven Einfluss, jedoch ohne eine Effektstärke oder die hierfür nötigen Daten anzugeben. Die verbleibenden drei Studien geben alle Effektstärken an und finden tendenziell hohe Effekte. Ein außergewöhnlich großer Effekt (Cohens'd >2) zeigt sich bei Bello et al. (2016), obgleich ein hoher Teil dieses Effektes auf das unzureichende Studiendesign fallen dürfte⁸.

⁸ Die Erhebung findet in einer Sekundarstufe statt. Verglichen werden ein interaktives Planspiel und eine Vorlesung, wie es scheint ganz ohne weitere Unterrichtsmaterialien.

Die meisten Studien kommen zu einer positiven Einschätzung hinsichtlich des Lernens mit Planspielen. Dieser Trend zeigt sich sowohl für Befragungsstudien als auch für Erhebungen, die den Lernerfolg anhand externer Daten beurteilen.

Fazit und Diskussion

Die Arbeit an diesem Review förderte eine große Fülle an unterschiedlichen Forschungsarbeiten zum Lernen mit Planspielen zutage. Dies trifft sowohl auf die thematische Breite der beforschten Planspiele zu als auch auf die Kulturen, in denen die Forschungsarbeiten entstanden sind. Sie stammen u.a. aus den Politikwissenschaften, der Pädagogik, den Wirtschaftswissenschaften und der Informatik. Dies zeigt das thematische Spektrum, in dem mit Planspielen gelehrt und in dem auch an Planspielen geforscht wird. In den 58 ausgewerteten Artikeln finden sich Arbeiten aus Afrika, Asien, Europa und Nordamerika. Es wäre lohnend einen vergleichenden Blick auf den Einsatz und die Erforschung von Planspielen zu werfen und herauszufinden, inwiefern sich Simulationen, die Spielbegleitung und Forschungsansätze sowohl kulturell als auch hinsichtlich der Fachbereiche unterscheiden.

Die große Breite der Forschung bringt auch Schwierigkeiten mit sich. Es scheint international unklar zu sein, was ein Planspiel eigentlich ist. Während im deutschsprachigen Raum häufig die Definition von Kriz (2011) aufgegriffen wird, scheint sich im englischsprachigen Raum keine Definition durchgesetzt zu haben. So werden unter dem Begriff „simulation game“ Computerspiele, Lernspiele und Planspiele gleichermaßen diskutiert. Die international ausgerichtete Forschung steht daher vor der Herausforderung mit schwammigen Begriffen arbeiten zu müssen und im Einzelfall zu entscheiden, inwiefern ein simulation game einem Planspiel entspricht. Aus den 58 Forschungsarbeiten, die in diesem Literature Review behandelt wurden, konnten zentrale Erkenntnisse zu Gelingensbedingungen extrahiert werden. Diese Erkenntnisse wurden strukturiert nach Planspiel, Spielleitung, Studierende und Lehrerfolg beschrieben. Es wurde deutlich, dass Planspiele hinsichtlich Schwierigkeit, Lernfortschritt und sprachlich/kulturellem Verständnis an ihre Zielgruppe anpassbar sein sollten. Authentisch sollten Planspiele insbesondere in Bezug auf Informationen/Daten und Wirkungsverläufe sein. Die Spiele sollten so gestaltet sein, dass sie leicht

handhabbar, inhaltlich aber tiefgründig, unvorhersehbar und komplex (aber dennoch an das Niveau der Lernenden anpassbar) sind. Die Bedeutung der Lehrenden in Planspielen ist zentral. Sie moderieren zwischen Simulation und Lernenden. Erst durch Einführung, Anpassung, Erklärung, Feedback und schlussendlich die Auswertung der Simulation entsteht für Lernende ein Mehrwert. Hinsichtlich der Studierenden wurden insbesondere das persönliche Engagement und die Art der Teamkommunikation als relevante Einflussfaktoren identifiziert. Viele der referierten Studien weisen darauf hin, dass sich Planspiele gut als Lernmethode eignen. Bemerkenswert ist die Breite der Lerngegenstände, die mit Planspielen adressiert werden.

Die Literaturstudie weist auf die hohe Komplexität planspielbasierter Lehre hin. Erst durch das Zusammenwirken von Planspiel, Spielbegleitung (Briefing, Debriefing, Didaktik) und der Bereitschaft der Studierenden, sich auf die Methode einzulassen, entfaltet sich das Potential von Planspielen für die Lehre: Das problem- und erfahrungsorientierte Erlernen von komplexen, systemischen Zusammenhängen. An zwei Beispielen kann gezeigt werden, dass die Einflussfaktoren nicht unabhängig voneinander agieren, sondern stark aufeinander bezogen sind: Sowohl unter dem Einflussfaktor Planspiel als auch unter dem Einflussfaktor Spielleitung wurde das Thema Feedback an Lernende als wichtige Kategorie identifiziert. Der Anspruch, Lernende im Verlauf des Lernprozesses mit Feedback zu begleiten, kommt zugleich aus zwei Richtungen und muss in der Praxis aufeinander abgestimmt werden. Je nachdem wie detailliert das Feedback des Planspiels ist, können Lehrende ihr Feedback knapper oder ausführlicher ausfallen lassen. Bei komplexen Planspielen mit hoher Informationsdichte könnten Studierende das vom Planspiel gegebene Feedback übersehen (oder missinterpretieren) und Lehrende müssen in ihrem Feedback das Feedback des Planspiels aufgreifen. Jedes Feedback, von Seiten des Planspiels oder der Lehrenden, muss von Studierenden aufgegriffen und verarbeitet werden, was wiederum auf das Engagement der Lernenden verweist. Ähnliches gilt für das Debriefing von Planspielen, das in dieser Arbeit unter dem Einflussfaktor Spielleitung behandelt wird (siehe oben). Gleich zwei Studien betonen hierbei auch die Rolle der Studierenden (Bereitschaft zum Teilen von Informationen und Einbringen evtl. nicht intendierter Lerninhalte). Auch dies weist auf den multifaktoriellen Charakter der Lehr-Lernforschung (mit Planspielen) hin. Es ist das maßvolle Zusammenspiel der Faktoren, das zum Gelingen planspielbasierter

Lehrveranstaltungen beiträgt. Isoliert determinierende Faktoren zu benennen, würde der Komplexität der Methode nicht gerecht und entspricht auch nicht der Idee des multifaktoriellen Modells der Lehrveranstaltungsqualität (Rindermann, 2009), in dem ebenfalls das Zusammenspiel der Faktoren betont wird.

Die Arbeit an diesem Review zeigt auch Schwächen und Lücken in der Lehr-Lernforschung mit Planspielen auf. Lernerfolge werden in den ausgewerteten Studien zumeist durch Befragung, einfache Wissenstests oder Spielergebnisse erhoben. Als problem- und erfahrungsorientierte Lehr-Lernmethode enthalten Planspiele jedoch tiefere Lernpotentiale, als z.B. durch Wissenstest abgefragt werden können. Wünschenswert wäre daher die Entwicklung von Forschungsdesigns, die das Lernpotential der Methode sichtbar machen können (z.B. Verständnis komplexer Wirkungsgefüge, systemischer und sozialer Zusammenhänge, Entscheidungskompetenz). Ansätze finden sich bei Plamunen et al. (2013, S. 854), die das systemische Verständnis von Studierenden anhand von Concept Maps erheben oder bei Pasi & Giroux., die die Messung des Lernerfolgs anhand des Spielerfolgs kritisieren und alternativen entwickeln (Pasi & Giroux, 2011, S. 1248). Vielversprechend scheinen auch qualitative Forschungsansätze, mit denen Reflexionsaufgaben von Studierenden ausgewertet werden (Ben-Zvi & Carton, 2008). Wünschenswert für die zukünftige Forschung sind triangulative Ansätze, die der Komplexität des Lernprozesses anhand ihrer methodischen Vielfalt eher entsprechen.

Eine deutliche Forschungslücke zeigt sich beim Einflussfaktor Rahmenbedingungen. Keine der ausgewerteten Studien nimmt den Einfluss von Rahmenbedingungen (z.B. Kursgröße, räumliche Ausstattung, Freiwilligkeit, Verbindung mit Prüfung) auf die Lehre mit Planspielen in den Blick. Zukünftige Forschungsvorhaben sollten hier ansetzen. Insbesondere für die curriculare und didaktische Weiterentwicklung planspielbasierter Lehre wären Forschungsergebnisse in diesem Bereich sehr lohnend. Dass die Rolle der Rahmenbedingungen in der Forschung zu planspielbasierten Lehrveranstaltungen nicht in den Blick genommen wird, erklärt sich teilweise durch eine weitere Schwäche der Planspielforschung. Sie ist sehr stark auf einzelne Beispiele konzentriert. Um unterschiedliche Rahmenbedingungen miteinander vergleichen zu können, bräuchte es jedoch unterschiedliche Beispiele bei gleichem Forschungsdesign (z.B. mit und ohne Prüfung, räumliche Ausstattung A vs. B, deutlich unterschiedliche Studierendenzahlen). Die Forschung konzentriert sich

jedoch bislang sehr stark auf spezifische Aspekte einzelner Planspiele (Erlernen von x mit Planspiel y). In nur vier der hier ausgewerteten Studien werden beispielsweise Daten von mehr als einem Planspiel erhoben. Wünschenswert wären Forschungsdesigns, mit denen es möglich ist, Vergleiche zwischen unterschiedlichen Planspielen und unterschiedlichen Arten der Durchführung zu ziehen und so strukturelle Unterschiede und Zusammenhänge planspielbasierter Lehre aufzuzeigen, die über das konkrete Beispiel hinausgehen.

Limitationen

Ein Systematic Review kann nur diejenigen Inhalte behandeln, die in der ausgewerteten Literatur zutage gefördert werden. Um die Arbeit in einem zu bewältigenden Umfang zu halten, konzentriert sich diese Arbeit auf zwei Datenbanken und verzichtete auf das systematische Durchsuchen von Literaturverzeichnissen. Eine erweiterte Literatursuche hätte eventuell weitere Aspekte berücksichtigen können. Die vielen signifikanten Ergebnisse (insbesondere beim Lernerfolg) könnten auf einen publication bias in der ausgewählten Literatur hinweisen (vgl. Newman & Gough, 2020, S. 8), wonach positive Ergebnisse überrepräsentiert sind, weil nicht signifikante oder negative Ergebnisse weniger publiziert werden bzw. schwerer zu finden sind. Die Qualität der Studien in der Forschung rund um Planspiele zeigt sich als sehr unterschiedlich. Neben komplexen Forschungsdesigns (exemplarisch: Oberle & Leunig, 2016; Pasin & Giroux, 2011) finden sich auch leicht kritisierbare Ansätze (exemplarisch: Bello et al., 2016). Auch wenn die Aussagen dieser Arbeit zumeist durch mehrere Ergebnisse gestützt werden, sollten dies bei der Interpretation der Ergebnisse doch berücksichtigt werden.

Literaturverzeichnis

Abdullah, N. L., Hanafiah, M. H. & Hashim, N. A. (2013). Developing Creative Teaching Module: Business Simulation in Teaching Strategic Management. *International Education Studies*, 6(6), 95–107.

- Alas, R., Kross, P., Liivat, A. & Saar, J. (2018). Success Factors for Teams in Business Game Dynama. *Problems and Perspectives in Management*, 16(1), 309–319. [https://doi.org/10.21511/ppm.16\(1\).2018.30](https://doi.org/10.21511/ppm.16(1).2018.30)
- Auman, C. (2011). Using Simulation Games to Increase Student and Instructor Engagement. *College Teaching*, 59(4), 154-161. <https://doi.org/10.1080/87567555.2011.602134>
- Bachen, C. M., Hernandez-Ramos, P. F. & Raphael, C. (2012). Simulating REAL LIVES: Promoting Global Empathy and Interest in Learning through Simulation Games. *Simulation & Gaming*, 43(4), 437–460. <https://doi.org/10.1177/1046878111432108>
- Bello, S., Ibi, M. B. & Bukar, I. B. (2016). Effect of Simulation Techniques and Lecture Method on Students' Academic Performance in Mafoni Day Secondary School Maiduguri, Borno State, Nigeria. *Journal of Education and Practice*, 7(23), 113–117.
- Ben-Zvi, T. & Carton, T. C. (2008). Applying Bloom's Revised Taxonomy in Business Games. *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL conference*, 35, 265–272. <https://absel-ojs-ttu.tdl.org/absel/issue/view/35>
- Beranic, T. & Hericko, M. (2019). Introducing ERP Concepts to IT Students Using an Experiential Learning Approach with an Emphasis on Reflection. *Sustainability*, 11(18), 1–17. <https://doi.org/10.3390/su11184992>
- Bioglio, L., Capecchi, S., Peiretti, F., Sayed, D., Torasso, A. & Pensa, R. G. (2019). A Social Network Simulation Game to Raise Awareness of Privacy among School Children. *IEEE Transactions on Learning Technologies*, 12(4), 456–469. <https://doi.org/10.1109/TLT.2018.2881193>
- Brazhkin, V. & Zimmerman, H. (2019). Students' Perceptions of Learning in an Online Multiround Business Simulation Game: What Can We Learn from Them? *Decision Sciences Journal of Innovative Education*, 17(4), 363–386. <https://doi.org/10.1111/dsji.12189>
- Brennan, R. & Vos, L. (2013). Effects of Participation in a Simulation Game on Marketing Students' Numeracy and Financial Skills. *Journal of Marketing Education*, 35(3), 259–270. <https://doi.org/10.1177/0273475313482928>
- Buil, I., Catalán, S. & Martínez, E. (2018). Exploring Students' Flow Experiences in Business Simulation Games. *Journal of Computer Assisted Learning*, 34(2), 183–192. <https://doi.org/10.1111/jcal.12237>
- Buil, I., Catalán, S. & Martínez, E. (2020). Engagement in Business Simulation Games: A Self-System Model of Motivational Development. *British Journal of Educational Technology*, 51(1), 297–311. <https://doi.org/10.1111/bjet.12762>
- Chen, L., Keys, A. & Gaber, D. (2015). How Does ERPsim Influence Students' Perceived Learning Outcomes in an Information Systems Course? An Empirical Study. *Journal of Information Systems Education*, 26(2), 135–146.
- Corbeil, P. & Laveault, D. (2011). Validity of a Simulation Game as a Method for History Teaching. *Simulation & Gaming*, 42(4), 462–475. <https://doi.org/10.1177/1046878108325451>
- Cronan, T. P., Leger, P.-M., Robert, J., Babin, G. & Charland, P. (2012). Comparing Objective Measures and Perceptions of Cognitive Learning in an ERP

- Simulation Game: A Research Note. *Simulation & Gaming*, 43(4), 461–480.
<https://doi.org/10.1177/1046878111433783>
- deNoyelles, A. & Raider-Roth, M. (2016). Being an “Agent Provocateur”: Utilising Online Spaces for Teacher Professional Development in Virtual Simulation Games. *Technology, Pedagogy and Education*, 25(3), 337–353.
<https://doi.org/10.1080/1475939X.2015.1049652>
- Dick, G. N. & Akbulut, A. Y. (2020). Innovative Use of the ERPsim Game in a Management Decision Making Class: An Empirical Study. *Journal of Information Technology Education: Research*, 19, 615–637.
- Dube, E. A. (2015). Using Games in Business Education: An Evaluation Experiment Comparing Games to Other Selected Methods in Teaching Sustainable Development Concepts. *International Journal for Business Education*, 155(1).
<https://doi.org/10.30707/IJBE155.1.1648133093.381331>
- Duke, R. D. (1980). A Paradigm for Game Design. *Simulation & Games*, 11(3), 364–377. <https://doi.org/10.1177/104687818001100308>
- Eränpalo, T. (2014). Exploring Young People’s Civic Identities through Gamification: A Case Study of Finnish, Swedish and Norwegian Adolescents Playing a Social Simulation Game. *Citizenship, Social and Economics Education*, 13(2), 104–120. <https://doi.org/10.2304/csee.2014.13.2.104>
- Geithner, S. & Menzel, D. (2016). Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. *Simulation & Gaming*, 47(2), 228–256. <https://doi.org/10.1177/1046878115624312>
- Gough, D., Oliver, S. & Thomas, J. (2017). Introducing systematic reviews. In D. Gough, S. Oliver & J. Thomas (Hrsg.), *An introduction to systematic reviews* (2. Aufl., S. 1–18). Sage.
- Gough, D., Thomas, J. & Oliver, S. (2012). Clarifying Differences between Review Designs and Methods. *Systematic Reviews*, 28(1), 1–9.
<https://doi.org/10.1186/2046-4053-1-28>
- Hamzeh, F., Theokaris, C., Rouhana, C. & Abbas, Y. (2017). Application of Hands-On Simulation Games to Improve Classroom Experience. *European Journal of Engineering Education*, 42(5), 471–481.
<https://doi.org/10.1080/03043797.2016.1190688>
- Herzig, M. (2019). Fame and Fortune: Developing a Simulation Game for the Music Industry Classroom. *Journal of the Scholarship of Teaching and Learning*, 19(5), 105–122. <https://doi.org/10.14434/josotl.v19i5.24276>
- Hilliard, I. (2014). People Power – Computer Games in the Classroom. *Higher Learning Research Communications*, 4(1).
<https://doi.org/10.18870/hlrc.v4i1.200>
- Hwang, M. (2019). ERP Simulation Games in Asynchronous Online Classes. *e-Journal of Business Education and Scholarship of Teaching*, 13(3), 37–45.
- Hwang, M. I. (2018). Relationship between Teamwork and Team Performance: Experiences from an ERPsim Competition. *Journal of Information Systems Education*, 29(3), 157–167.
- Imhof, M. & Starker, U. (2020). Das psychologische Planspiel als didaktisches Szenario – Anwendung und Nutzen. In M. Krämer, J. Zumbach & I. Deibl

- (Hrsg.), Materialien aus der Sektion Aus-, Fort- und Weiterbildung in Psychologie (AFW) im Berufsverband Deutscher Psychologinnen und Psychologen e.V. (BDP): Bd. 16. Psychologiedidaktik und Evaluation XIII. Shaker Verlag. <https://doi.org/10.23668/psycharchives.4252>
- Kriz, W. C. (2009). Planspiel. In S. Kühl, P. Strodholz & A. Taffertshofer (Hrsg.), *Handbuch Methoden der Organisationsforschung* (S. 558–580). VS Verlag für Sozialwissenschaften (GWV).
- Kriz, W. C. (2011). Qualitätskriterien von Planspielanwendungen. In S. Hitzler (Hrsg.), *ZMS-Schriftenreihe: Bd. 2. Planspiele - Qualität und Innovation: Neue Ansätze aus Theorie und Praxis* (1. Aufl., S. 11–38). Books on Demand.
- Kwon, S., Lara, M., Enfield, J. & Frick, T. (2013). Design and Evaluation of a Prompting Instrument to Support Learning within the Diffusion Simulation Game. *Journal of Educational Technology Systems*, 41(3), 231–253. <https://doi.org/10.2190/ET.41.3.c>
- Leigh, E. & Spindler, L. (2005). Congruent Facilitation of Simulations and Games. In R. Shiratori, K. Arai & F. Kato (Hrsg.), *Gaming, Simulations, and Society* (S. 189–198). Springer-Verlag. https://doi.org/10.1007/4-431-26797-2_20
- Lin, Y.-L. & Tu, Y.-Z. (2012). The Values of College Students in Business Simulation Game: A Means-End Chain Approach. *Computers & Education*, 58(4), 1160–1170. <https://doi.org/10.1016/j.compedu.2011.12.005>
- Mandl, H. (2004). On the Path to a New Learning Culture – the Contribution of Web-Based Business Games. In T. Eberle (Hrsg.), *Bridging the Gap: Transforming Knowledge into Action through Gaming and Simulation: Proceedings of the 35th Conference of the International Simulation and Gaming Association, Munich, 2004* (S. 163–174).
- Matute-Vallejo, J. & Melero-Polo, I. (2019). Understanding Online Business Simulation Games: The Role of Flow Experience, Perceived Enjoyment and Personal Innovativeness. *Australasian Journal of Educational Technology*, 35(3), 71-85. <https://doi.org/10.14742/ajet.3862>
- Mayer, I., Warmelink, H. & Bekebrede, G. (2013). Learning in a Game-Based Virtual Environment: A Comparative Evaluation in Higher Education. *European Journal of Engineering Education*, 38(1), 85-106. <https://doi.org/10.1080/03043797.2012.742872>
- Misfeldt, M. (2015). Scenario Based Education as a Framework for Understanding Students Engagement and Learning in a Project Management Simulation Game. *Electronic Journal of e-Learning*, 13(3), 181–191.
- Mohd Radzi, S. H., Tan, W. H. & Yusoff, A. (2019). Shipping Management Simulation Game for Teaching and Learning in Higher Education: A Quasi-Experimental Study. *Malaysian Journal of Learning and Instruction*, 16(2), 155–186. <http://e-journal.uum.edu.my/index.php/mjli/article/view/mjli2019.16.2.6>
- Mohsen, K., Abdollahi, S. & Omar, S. (2019). Evaluating the Educational Value of Simulation Games: Learners' Perspective. *Innovations in Education and Teaching International*, 56(4), 517-528. <https://doi.org/10.1080/14703297.2018.1515646>

- Monk, E. F. & Lycett, M. (2016). Measuring Business Process Learning with Enterprise Resource Planning Systems to Improve the Value of Education. *Education and Information Technologies*, 21(4), 747-768.
<https://doi.org/10.1007/s10639-014-9352-6>
- Newman, M. & Gough, D. (2020). Systematic Reviews in Educational Research: Methodology, Perspectives and Application. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond & K. Buntins (Hrsg.), *Systematic Reviews in Educational Research* (S. 3–22). Springer Fachmedien Wiesbaden.
https://doi.org/10.1007/978-3-658-27602-7_1
- Ney, M., Gonçalves, C. & Balacheff, N. (2014). Design Heuristics for Authentic Simulation-Based Learning Games. *IEEE Transactions on Learning Technologies*, 7(2), 132–141. <https://doi.org/10.1109/TLT.2014.2316161>
- Nguyen, T. N. (2015). Motivational Effect of Web-Based Simulation Game in Teaching Operations Management. *Journal of Education and Training Studies*, 3(2), 9–15. <https://doi.org/10.11114/jets.v3i2.565>
- Norris, D. R. (2013). Beat the Bourgeoisie: A Social Class Inequality and Mobility Simulation Game. *Teaching Sociology*, 41(4), 334-345.
<https://doi.org/10.1177/0092055X13490751>
- Nussbaum, E. M., Owens, M. C., Sinatra, G. M., Rehmat, A. P., Cordova, J. R., Ahmad, S., Harris, Fred C., Jr. & Dascalu, S. M. (2015). Losing the Lake: Simulations to Promote Gains in Student Knowledge and Interest about Climate Change. *International Journal of Environmental and Science Education*, 10(6), 789–811.
- Oberle, M. & Leunig, J. (2016). Simulation Games on the European Union in Civics: Effects on Secondary School Pupils' Political Competence. *Citizenship, Social and Economics Education*, 15(3), 227–243.
<https://doi.org/10.1177/2047173416689794>
- Palmunen, L.-M., Pelto, E., Paalumäki, A. & Lainema, T. (2013). Formation of Novice Business Students' Mental Models through Simulation Gaming. *Simulation & Gaming*, 44(6), 846–868. <https://doi.org/10.1177/1046878113513532>
- Pashler, H., McDaniell, M., Rohrer, D. & Bjork, R. (2008). Learning Styles: Concepts and Evidence. *Psychological Science in the Public Interest*, 9(3), 105–119. <https://doi.org/10.1111/j.1539-6053.2009.01038.x>
- Pasin, F. & Giroux, H. (2011). The Impact of a Simulation Game on Operations Management Education. *Computers & Education*, 57(1), 1240–1254.
<https://doi.org/10.1016/j.compedu.2010.12.006>
- Piercy, N., Brandon-Jones, A., Brandon-Jones, E. & Campbell, C. (2012). Examining the Effectiveness of Experiential Teaching Methods in Small and Large OM Modules. *International Journal of Operations and Production Management*, 32(12), 1473–1492. <https://doi.org/10.1108/01443571211284205>
- Prihadi, K., Cheow, D. Z. Y., Yong, J. H. E. & Sundrasagran, M. (2018). Improving Resilience and Self-Esteem among University Students with Entrepreneurship Simulation Board Game. *International Journal of Evaluation and Research in Education*, 7(1), 48–56. <https://doi.org/10.11591/ijere.v7i1.11406>

- Richards, B. & Camuso, L. (2015). Cultural Capital in the Classroom: The Significance of Debriefing as a Pedagogical Tool in Simulation-Based Learning. *International Journal of Teaching and Learning in Higher Education*, 27(1), 94–103.
- Rindermann, H. (2009). *Lehrevaluation: Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen mit einem Beitrag zur Evaluation computerbasierten Unterrichts* (2. Aufl.). *Psychologie: Bd. 42*. Empirische Pädagogik e.V.
- Sandberg, J. A. C., Wielinga, B. J. & Christoph, L. H. (2012). The Role of Prescriptive Models in Learning. *Computers & Education*, 59(2), 839-854. <https://doi.org/10.1016/j.compedu.2011.11.021>
- Schäfer, E. (2017). *Lebenslanges Lernen*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-50422-2>
- Scholkmann, A., Siemon, J., Boom, K.-D. & Knigge, M. (2017). Lernzeitnutzung im Planspielunterricht. Eine Analyse des Einflusses kognitiver Fähigkeiten, Zielorientierungen und Charakteristika von Lernpartnern anhand von Videodaten. *Zeitschrift für Erziehungswissenschaft*, 20(4), 651–669. <https://doi.org/10.1007/s11618-017-0736-4>
- Schwägele, S. (2015). *Planspiel - Lernen - Lerntransfer: Eine subjektorientierte Analyse von Einflussfaktoren*. Dissertation. ZMS-Schriftenreihe: Band 7.
- Schwägele, S., Zürn, B., Lukosch, H. K. & Freese, M. (2021). Design of an Impulse-Debriefing-Spiral for Simulation Game Facilitation. *Simulation & Gaming*, 52(3), 364–365. <https://doi.org/10.1177/10468781211006752>
- Seethamraju, R. (2011). Enhancing Student Learning of Enterprise Integration and Business Process Orientation through an ERP Business Simulation Game. *Journal of Information Systems Education*, 22(1), 19–29.
- Siewiorek, A., Gegenfurtner, A., Lainema, T., Saarinen, E. & Lehtinen, E. (2013). The Effects of Computer-Simulation Game Training on Participants' Opinions on Leadership Styles. *British Journal of Educational Technology*, 44(6), 1012–1035. <https://doi.org/10.1111/bjet.12084>
- Siewiorek, A., Saarinen, E., Lainema, T. & Lehtinen, E. (2012). Learning Leadership Skills in a Simulated Business Environment. *Computers & Education*, 58(1), 121–135. <https://doi.org/10.1016/j.compedu.2011.08.016>
- Tao, Y.-H., Yeh, C. R. & Hung, K. C. (2012). Effects of the Heterogeneity of Game Complexity and User Population in Learning Performance of Business Simulation Games. *Computers & Education*, 59(4), 1350–1360. <https://doi.org/10.1016/j.compedu.2012.06.003>
- Treen, E., Atanasova, C., Pitt, L. & Johnson, M. (2016). Evidence from a Large Sample on the Effects of Group Size and Decision-Making Time on Performance in a Marketing Simulation Game. *Journal of Marketing Education*, 38(2), 130-137. <https://doi.org/10.1177/0273475316653433>
- van den Bossche, P., Gijssels, W., Segers, M., Woltjer, G. & Kirschner, P. (2011). Team Learning: Building Shared Mental Models. *Instructional Science*, 39(3), 283–301.
- Williams, D. (2015). The Impact of SimVenture on the Development of Entrepreneurial Skills in Management Students. *Industry and Higher Education*, 29(5), 379–395. <https://doi.org/10.5367/ihe.2015.0270>

Zulfiqar, S., Sarwar, B., Aziz, S., Ejaz Chandia, K. & Khan, M. K. (2019). An Analysis of Influence of Business Simulation Games on Business School Students' Attitude and Intention toward Entrepreneurial Activities. *Journal of Educational Computing Research*, 57(1), 106–130.
<https://doi.org/10.1177/0735633117746746>

5.2 Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre

Abstract

Während der pandemiebedingten Distanzlehre mussten auch Planspiele auf ein Onlineformat umgestellt werden. Aufgrund der Komplexität der Methode war die Herausforderung hier besonders hoch. Der Beitrag analysiert anhand von Evaluationsdaten aus 124 planspielbasierten Lehrveranstaltungen welche Gemeinsamkeiten und Unterschiede beim Einsatz von synchronen Präsenz- und Onlineplanspielen in der Hochschullehre bestehen. Der Vergleich zeigt in mehreren Dimensionen (z.B. Lehrendenverhalten, Mitarbeit der Studierenden, Teamarbeit) zwar signifikante, allerdings eher kleine Unterschiede. Deutlichere Unterschiede zeigt der Vergleich Online/Präsenz lediglich bei haptischen Planspielen. Die Autoren folgern, dass die Übertragung von Planspielen in die Onlinelehre insgesamt als gelungen angesehen werden kann.

DOI: [10.21240/zfhe/SH-PS/19](https://doi.org/10.21240/zfhe/SH-PS/19)

Autoren: Tobias Alf, Friedrich Trautwein

Einleitung

Zu Beginn der Coronapandemie standen Lehrende vor der Herausforderung in kürzester Zeit auf Distanzlehre umzustellen. Angesichts der Komplexität von Planspielen als Lehr-Lernmethode stellte dies eine besondere Herausforderung dar. So zeichnen sich Planspiele durch ein hohes Maß an Interaktion aus, erfordern die Zusammenarbeit in Gruppen, basieren auf spezifischem Material und enthalten oft haptische Elemente. Im Vergleich zu traditionellen Lehrveranstaltungen erschwert dies die Übertragung in ein Distanzformat enorm. Unter Verwendung eines am Zentrum für Managementsimulation (ZMS) der DHBW Stuttgart⁹ entwickelten Fragebogens zur Evaluation von Planspiellehrveranstaltungen werden Vergleiche (t-Test und ANOVA) zwischen Präsenz- und Onlineplanspielen gezogen, um künftige Entscheidungen zum Einsatz von Planspielen evidenzbasiert treffen zu können. In den Kapiteln zwei und drei geben wir einen Überblick über die Rolle von Planspielen in der Hochschullehre und beschreiben, was genau in diesem Beitrag unter Onlinelehre verstanden wird. Kapitel 4 beschreibt die konkreten Forschungsfragen, gefolgt von einer Vorstellung des Fragebogens und der Datenerhebung (Kapitel 5). In Kapitel 6 werden die Ergebnisse je Forschungsfrage dargestellt, bevor sie in Kapitel 7 diskutiert und bewertet werden. Der Artikel schließt mit einigen Restriktionen.

Hochschullehre in Präsenz und Online

Schon vor der Coronapandemie wurde an Hochschulen auf Distanz gelehrt und gelernt – und dazu geforscht. Tendenziell ist der Lernerfolg in Distance Education Programmen geringer als in Präsenzprogrammen (Simons et al., 2020, S. 278; Xu & Xu, 2020, S. 376). Die diskutierten Gründe hierfür sind vielfältig: Teilnehmende in Distance Learning Programmen sind tendenziell älter und haben daher neben dem Studium familiäre und berufliche Verpflichtungen (Simons et al. 2020). Zudem scheint ein „lack of interpersonal connection“, sowohl zu Lehrenden als auch zu Peers, das Studium zu erschweren (Xu & Xu, 2020, S. 382). Auch die höheren

⁹ Das ZMS ist ein Lehr- und Forschungszentrum, das sich in Lehre und Forschung mit Planspielen beschäftigt (vgl. hierzu <https://zms.dhbw-stuttgart.de/>).

Anforderungen an eigenständiges Lernen („self directed learning“) werden als Grund diskutiert (ebd.; Broadbent & Poon, 2015; Torun, 2021). Hierzu passt, dass als Prädikatoren für erfolgreiches Distance Learning gerade „self directed learning“ (Torun, 2021, S. 191) und „peer learning“ (Broadbent & Poon, 2015, S. 12) diskutiert werden. Da biografische Gründe (Familie, Beruf) für die coronabedingte Onlinelehre nicht ausschlaggebend sein dürften, können die soziale Einbindung in die Lehrveranstaltung und self directed learning als die kritischen Punkte angesehen werden.

Vermutlich bedingt durch die Dringlichkeit der pandemiebedingten Onlinelehre, sind sehr unscharfe Diskussionen und Begriffe zur Beschreibung bestimmter Konstellationen des digital gestützten Lernens entstanden¹⁰. Aber auch in den älteren, oben genannten Studien wird häufig nicht deutlich ausgeführt was Onlinelehre genau meint. Die begriffliche Vielfalt für gleiche Sachverhalte bzw. die begriffliche Gleichheit für verschiedene Sachverhalte erschwert die Diskussion erheblich. Im Folgenden beschreiben wir daher explizit, was im Rahmen dieser Arbeit unter Präsenz- und Onlinelehre verstanden wird. Hierzu beziehen wir uns auf die zweidimensionale Kategorisierung von Lehrformen nach Entner et al. (2021) (Tabelle 1), in der einerseits die Unterscheidung physisch/virtuell und andererseits die Unterscheidung synchron/asynchron vorgenommen wird. Es ergeben sich somit vier Felder, zur Unterscheidung von Lehrformen:

Tabelle 1: Zweidimensionale Kategorisierung von Lehrformen, eigene Darstellung nach ENTNER et al. (2021)

	Synchron	Asynchron
Physischer Raum (Offline)	z.B. Vorlesungen im Hörsaal, Tutorien im Kleingruppenraum	z.B. Vorlesungsskript, Aufgabenblätter
Virtueller Raum (Online)	z.B. Vorlesungen im Live-Stream, Chat zu Übungsaufgaben	z.B. Vorlesungsaufzeichnung, Diskussions-Forum zu Übungen

¹⁰ Zum Beispiel wird der Begriff „hybride Lehre“ einerseits als gleichzeitiges Lernen von anwesenden Studierenden und per Videokonferenz zugeschalteten Studierenden verstanden (Fallmann et al., 2021; Rachbauer und Hanke, 2022), andererseits wird er mit Blended Learning gleichgesetzt (Entner et al., 2021; Reinmann, 2021).

Onlinelehre bezieht sich in diesem Beitrag auf synchrone Veranstaltungen. Es werden synchrone Veranstaltungen im physischen und virtuellen Raum miteinander verglichen (erste Spalte), während asynchrone Elemente keine Rolle spielen.

Planspiele in der Hochschullehre

In Planspielen werden Aspekte der Wirklichkeit modellhaft nachgebildet (Simulation). In dieser Simulation übernehmen Lernende Rollen und agieren als diese, unter Berücksichtigung von Spielregeln, um simulierte Probleme zu bearbeiten (Kriz, 2011). Planspiele enthalten somit Elemente aus Simulation, Regelspiel und Rollenspiel, ohne diesen aber ganz zugeordnet werden zu können (ebd.). Kennzeichnend für Planspiele ist die Abfolge von Briefing, Spielphase und Debriefing. In längeren rundenbasierten Planspielen wiederholt sich der Prozess aus Briefing, Spielphase und Debriefing mehrmals. Planspiele werden als problemorientierte (Mandl, 2004) und erfahrungsorientierte Lehr-Lernmethode diskutiert (Geithner & Menzel, 2016; Schwägele, 2015). Diese Einordnung entspricht dem charakteristischen Wechsel von Aktivität/Erfahrung einerseits und Reflexion der Erfahrung andererseits. Um Planspiele im größeren Kontext der Hochschullehre zu verorten, bietet sich das Modell der akademischen Lehre nach Reinmann (2016) an. In diesem wird akademische Lehre in drei Räume untergliedert. Während der Informationsraum die rezeptive Teilhabe an Wissen ermöglicht, eignen sich Studierende im Erprobungsraum Wissen aktiv an. Dort findet „produktives Lernen anhand komplexer Probleme“ statt, bei dem „konkrete Artefakte als Problemlösungen resultieren“ (ebd. 229). Der Explorationsraum („Lernen durch Forschung“) ermöglicht das Durchlaufen ganzer Forschungsprojekte, mit dem Potential neues Wissen zu generieren. Planspiele lassen sich in den Erprobungsraum des Modells der akademischen Lehre einordnen. Studierende werden aktiv und können sich erproben. Sie lernen produktiv entsprechend der im Planspiel entworfenen Szenarien, wobei auch Artefakte entstehen (Entscheidungen, Ausarbeitungen, Spielergebnisse etc.).

Forschungsfragen

Wie beschrieben, sind Planspiele eine Lehr-Lernmethode, in der Studierende zu aktivem (Übernahme von Rollen, Entscheidungen treffen, simulierte Probleme lösen) und produktivem (es entstehen Ergebnisse, Artefakte) Lernen angeregt werden. Zudem erfordern Planspiele ein hohes Maß an Kommunikationsmöglichkeiten. Vor diesem Hintergrund stellt sich die Frage, wie erfolgreich eine solch komplexe und interaktive Methode in der Onlinelehre umgesetzt werden kann. Konkret werden in diesem Beitrag die folgenden Forschungsfragen beantwortet:

- Forschungsfrage 1: Welche Unterschiede gibt es in der planspielbasierten Präsenz- und Onlinelehre im Hinblick auf Lehrende und Studierende?
- Forschungsfrage 2: Welche Unterschiede gibt es in der planspielbasierten Präsenz- und Onlinelehre hinsichtlich der Bewertung des Planspiels?
- Forschungsfrage 3: Welche Unterschiede gibt es in der planspielbasierten Präsenz- und Onlinelehre hinsichtlich Lernen und Zufriedenheit?

Empirische Erhebung

Grundlage für diese Studie ist ein Fragebogen zur Evaluation planspielbasierter Lehre, der am ZMS entwickelt wurde. Itemanalysen und explorative Faktorenanalysen deuten auf ein gutes Maß an Reliabilität und Validität hin (Trautwein & Alf, 2022), wobei dies in weiteren Studien bestätigt werden sollte. Der Bogen kombiniert Aspekte der allgemeinen Hochschullehre mit spezifischen Einflussfaktoren von Planspielen. Er orientiert sich an den von Rindermann (2003) beschriebenen Einflussfaktoren (Studierende, Lehrende, Rahmenbedingungen) und ergänzt das Planspiel und die Teamarbeit als relevante Faktoren für planspielbasierte Lehre. Alle Aspekte zusammen tragen zum Lernen und zur Zufriedenheit bei.

Tabelle 2: Überblick zu Items und Faktoren

Items	Skala	Alpha
Die Spielergebnisse lassen sich gut nachvollziehen. Die Spielergebnisse sind einleuchtend. Ich habe verstanden, wie das Planspiel funktioniert.	Planspiel: Verständlichkeit	.833
Das Planspiel ist praxisnah. Das Planspiel bildet die Realität gut ab.	Planspiel: Praxisbezug	.861
Die Planspielleitung kennt sich mit dem Planspiel gut aus. Die Veranstaltung war gut strukturiert.	Planspielleitung: Kompetenz	.782
Es bestand in ausreichendem Maß die Möglichkeit, sich mit der Planspielleitung zu beraten. Die Beratungen mit der Planspielleitung fanden in einer konstruktiven Atmosphäre statt. Die Unterstützung der Planspielleitung war fachlich hilfreich.	Planspielleitung: Kommunikation	.895
Es fiel uns leicht, in der Gruppe Entscheidungen zu treffen. Die zeitliche Abstimmung in der Gruppe hat gut funktioniert. Die Atmosphäre in der Gruppe war gut. Ich würde sagen, dass unser Team sehr gut organisiert war.	Team: Organisation und Atmosphäre	.851
Jeder hatte eine Funktion in der Gruppe. Es gab eine klare Aufgabenverteilung in der Gruppe. Ich wusste, was meine Aufgaben sind.	Team: Aufgabenverteilung	.814
In den Entscheidungsphasen (Arbeit in der Kleingruppe) habe ich mich aktiv beteiligt. Wenn andere meine Rolle betrachten, würden sie mich zu den aktiveren Teilnehmenden zählen.	Studierende: Engagement	.783
Ich habe beim Planspiel viel gelernt. Ich habe in dieser Veranstaltung etwas Sinnvolles und Wichtiges gelernt. Mein Verständnis für mein Studienfach hat sich durch die Veranstaltung weiterentwickelt. Alles in allem hat sich der Besuch dieser Veranstaltung für mich gelohnt. Das Planspiel war für mich sehr motivierend. Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?	Lernen und Zufriedenheit	.938

Entsprechend des Modells wurden für diese Studie aus 25 Items 8 Skalen gebildet (Tabelle 2). Die Fragen wurden auf einer sechsstufigen Likert-Skala beantwortet, die meist von "trifft gar nicht zu" (= 1) bis "trifft voll zu" (= 6) reicht. Zwei Skalen befassen sich mit Aspekten des Planspiels: Wie gut das Spiel von den Lernenden verstanden wird (Verständlichkeit) und wie gut die Simulation die reale Welt repräsentiert (Praxisbezug). Zwei Skalen messen lehrspezifische Aspekte der

Planspielleitung: Einerseits die wahrgenommene Kompetenz der Spielleitung und andererseits deren Kommunikation. Zwei weitere Skalen erheben Aspekte der Teamarbeit: Organisation und Atmosphäre fasst Items zusammen, die angeben, wie "leicht" oder "gut" die Zusammenarbeit im Team ist. Die Skala Aufgabenverteilung gibt Aufschluss über Funktionen und Aufgaben innerhalb des Teams. Die Skala Engagement gibt an, wie intensiv sich die Studierenden in das Spiel einbringen. In der Logik des Modells ist die Skala für Gesamtbewertung und Lernen die abhängige Variable, da sie für das Ziel der Lehre steht. Sie besteht aus drei Items, die sich auf das Lernen beziehen, und drei Items, die sich auf die allgemeine Zufriedenheit beziehen. Für alle Skalen zeigen sich Alpha Werte > 0.7 .

In die vorliegende Studie gehen Daten von 2119 Studierenden aus insgesamt 124 Planspielveranstaltungen mit 30 unterschiedlichen Planspielen ein. 43,4% der Studierenden identifizierten sich als männlich, 50,7% als weiblich, 0,9% als divers und weitere 5% machten hierzu keine Angabe. Knapp 25% der Befragten haben noch nie an einem Planspiel teilgenommen, alle anderen an mindestens einem. Die Daten wurden im Zeitraum von Mai 2021 bis August 2022 an der Fakultät Wirtschaft und Gesundheit der DHBW Stuttgart erhoben. Bei den eingesetzten Planspielen handelt es sich daher überwiegend um Wirtschaftsplanspiele¹¹. Hierunter sind 20 computergestützte Planspiele und 10 haptische Planspiele. Bei haptischen Planspielen erfolgt die Simulation auf einem Spielplan, auf dem Figuren und symbolhafte Materialien (Geld, Waren etc.) bewegt werden. Computergestützte Planspiele arbeiten mit Eingabemasken, in denen Entscheidungen übermittelt werden. Die Simulation erfolgt über programmierte Berechnungen. Insbesondere die Durchführung haptischer Planspiele in der Onlinelehre brachte große Herausforderungen mit sich. Zur Umsetzung haptischer Planspiele in der Onlinelehre, wurden vom ZMS anhand digitaler und kollaborativer Whiteboards Spieloberflächen entwickelt, um eigentlich haptische Vorgänge digital nachzubilden.

Zur Datenerhebung geben die Lehrenden am Ende eines Planspiels einen QR-Code aus, mit dem die Studierenden zum Fragebogen gelangen und ihn direkt ausfüllen. Auf diese Weise ergibt sich für evaluierte Veranstaltungen eine hohe

¹¹ Ein Überblick zu den am ZMS eingesetzten Planspielen findet sich unter: <https://zms.dhbw-stuttgart.de/das-zms/unsere-planspiele/>.

Rücklaufquote von ca. 80%, während für andere Veranstaltungen (Evaluation vergessen, keine Zeit mehr zur Verfügung) gar keine Daten vorliegen.

Alle evaluierten Veranstaltungen (Präsenz und Online) sind als Blockveranstaltungen konzipiert und können, je nach Planspiel, zwischen einem halben Tag und maximal drei Tagen dauern. In allen Fällen beginnt die Veranstaltung mit dem Briefing und endet mit einem Debriefing, ohne dass zeitlich vor- oder nachgelagerte Arbeitsformen vorgesehen sind (synchron). Während der ganzen Zeit sind Lehrende vor Ort bzw. im Onlinemeeting.

Ergebnisse

Die oben beschriebenen Forschungsfragen werden im Folgenden aufgegriffen und anhand der empirischen Daten analysiert und diskutiert. Zur besseren Lesbarkeit werden im Text nicht alle Zahlen berichtet. Eine detaillierte Übersicht der empirischen Daten findet sich in Tabelle 3, auf die jeweils verwiesen wird.

Unterschiede zwischen Präsenz- und Onlinelehre im Hinblick auf Lehrende und Studierende

Für Lehrende und Studierende stehen fünf Skalen zur Verfügung, wobei sich zwei auf Lehrende (Kompetenz und Kommunikation) und drei auf Studierende (Engagement, Team: Organisation und Atmosphäre sowie Team: Aufgabenverteilung) beziehen. Das Dozierendenverhalten wird in beiden Skalen signifikant positiver für die Onlinelehre eingeschätzt. Sowohl die Kompetenz der Dozierenden (5,43 vs. 5,23, $d=.221$) als auch deren Kommunikation (5,42 vs. 5,24, $d=.215$) werden online etwas höher eingeschätzt als in Präsenz (Tabelle 3). Ein Blick auf die Effektstärke Cohen's d^{12} zeigt jedoch, dass es sich um eher kleine Effekte handelt.

Auch für das Verhalten der Studierenden zeigen sich drei signifikante Ergebnisse, die jedoch nicht ganz einheitlich in eine Richtung weisen. Die Studierenden bewerten ihr Engagement in Onlineveranstaltungen höher als in

¹² Nach Cohen (1988, S. 25) handelt es sich bei ca. $d=0,2$ um einen kleinen Effekt, ab $d=0,5$ um einen mittleren Effekt und ab $d=0,8$ um einen großen Effekt.

Präsenzveranstaltungen (5,09 vs. 5,00, $d=.096$). Die Atmosphäre im Team (5,10 vs. 5,19; $d=.108$) und die Aufgabenorientierung bzw. Aufgabenverteilung im Team (4,47 vs. 4,69; $d=.179$) sind jedoch für Planspiele in Präsenz jeweils etwas höher (Tabelle 3). Dies scheint plausibel, da angenommen werden kann, dass die Bedingungen für gute Teamarbeit in Präsenz vorteilhafter sind. Auch hier zeigen die Effektstärken jedoch sehr kleine Werte, sodass die Unterschiede als gering einzuschätzen sind.

Tabelle 3: T-Tests: Vergleich von Online- und Präsenzlehre mit Planspielen

	alle Planspiele				computergestützte Planspiele				haptische Planspiele			
	N	Mean	Sig. (2-tailed)	Mean Difference	N	Mean	Sig. (2-tailed)	Mean Difference	N	Mean	Sig. (2-tailed)	Mean Difference
Dozierende	1136	5,43	0,001	0,19	925	5,46	0,037	0,10	211	5,27	0,038	0,16
Kompetenz	949	5,23			460	5,36			489	5,11		
Dozierende	1136	5,42	0,001	0,18	925	5,46	0,001	0,17	211	5,24	0,474	0,06
Kommunikation	949	5,24			460	5,29			489	5,18		
Team	1133	5,10	0,014	-0,09	922	5,09	0,259	-0,05	211	5,13	0,149	-0,10
Organisation u. Atmosphäre	947	5,19			460	5,14			487	5,23		
Team	1132	4,47	0,001	-0,22	921	4,41	0,041	-0,14	211	4,71	0,310	-0,10
Aufgabenverteilung	945	4,69			460	4,55			485	4,81		
Studierende	1129	5,09	0,030	0,10	918	5,11	0,386	0,05	211	5,03	0,277	0,09
Engagement	945	5,00			459	5,06			486	4,94		
Planspiel	1137	4,85	0,291	-0,04	926	4,81	0,001	0,19	211	5,04	0,087	-0,11
Verständlichkeit	951	4,89			460	4,61			491	5,16		
Planspiel	1136	4,32	0,657	0,02	925	4,38	0,035	0,13	211	4,06	0,001	-0,28
Praxisbezug	949	4,30			458	4,25			491	4,34		
Gesamtbew. und Lernen	1129	4,41	0,003	-0,15	920	4,49	0,479	-0,05	209	4,06	0,001	-0,52
	945	4,56			460	4,54			485	4,58		

Unterschiede zwischen Präsenz- und Onlinelehre im Hinblick auf die Bewertung des Planspiels

Auf den ersten Blick scheinen Studierende die Planspiele in der Präsenz- und Onlinelehre nicht unterschiedlich wahrzunehmen. Tabelle 3 zeigt für alle Planspiele keine signifikanten Unterschiede für die zwei Skalen zum Planspiel (Verständnis und Praxisbezug). Differenziert man jedoch weiter nach computergestützten Planspielen einerseits und haptischen Planspielen andererseits, so werden zumindest kleine Unterschiede deutlich: Verständnis (Präsenz 4,61 vs. Online 4,81) und Praxisbezug (Präsenz 4,25 vs. Online 4,38) des Planspiels werden in der Onlinelehre mit computergestützten Planspielen besser eingeschätzt. Bei haptischen Planspielen hingegen werden Praxisbezug (Präsenz 4,35 vs. Online 4,06) und tendenziell (nicht signifikant) auch das Verständnis (Präsenz 5,16 vs. Online 5,04) höher bewertet, wenn sie in Präsenz durchgeführt werden. Diese entgegengesetzten Effekte zwischen computerbasierten und haptischen Planspielen nivellieren sich gegenseitig, sodass kein Unterschied zwischen Präsenz- und Onlinelehre erkennbar ist, wenn man alle Planspiele betrachtet (Tabelle 3). Ein Blick auf die Effektstärken zeigt auch hier, dass die Unterschiede gering ausfallen. Trotz der herausfordernden Übertragung von haptischen Materialien in digitale Formate, waren auch haptische Planspiele in der Onlinelehre für Studierende gut nachvollziehbar und verständlich.

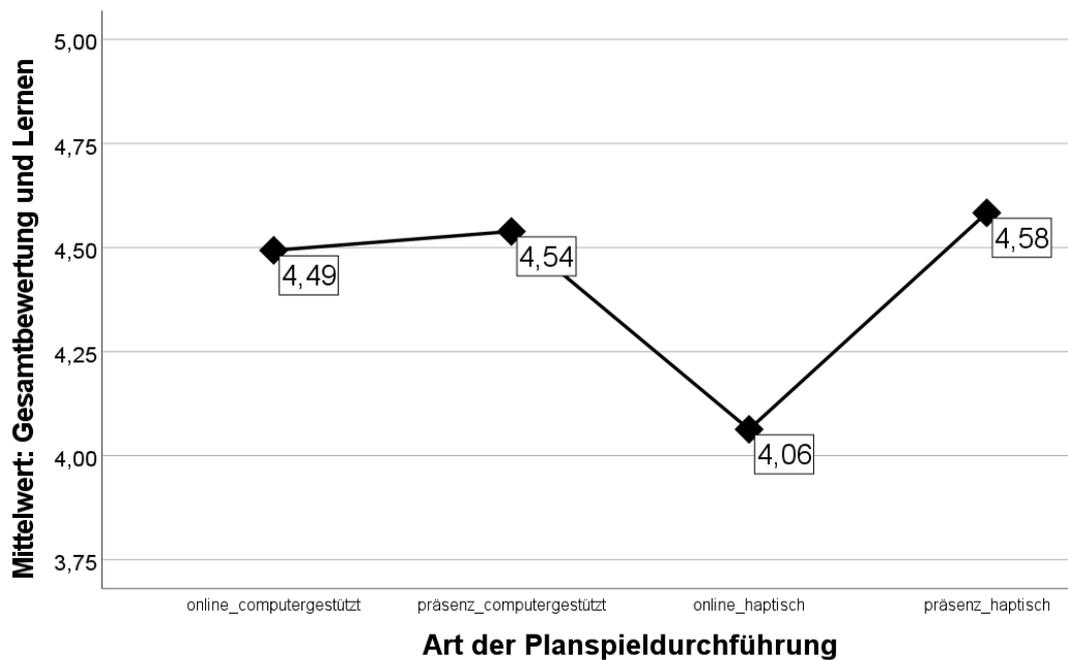
Unterschiede zwischen Präsenz- und Onlinelehre im Hinblick auf Lernen und Zufriedenheit

Hinsichtlich Lernen und Zufriedenheit schneiden Präsenzveranstaltungen etwas besser ab als Onlineplanspiele (4,56 vs. 4,41), auch hier mit geringer Effektstärke (vgl. Tabelle 3). Eine genauere Analyse mit Unterscheidung in computergestützte und haptische Planspiele in der Präsenz- und Onlinelehre zeigt jedoch, dass dieser Effekt auf die haptischen Planspiele in der Onlinelehre zurückgeführt werden kann (Tabelle 3). Zur weiteren Analyse wurden Gruppen mit vier unterschiedlichen Möglichkeiten der Planspieldurchführung gebildet:

- Onlinelehre mit computergestütztem Planspiel
- Präsenzlehre mit computergestütztem Planspiel
- Onlinelehre mit haptischem Planspiel
- Präsenzlehre mit haptischem Planspiel

Ein Vergleich der Mittelwerte dieser Gruppen bestätigt die vorherigen Analysen. Während die Variante Onlinelehre mit haptischem Planspiel für Lernen und Zufriedenheit auf einen Mittelwert von 4.1 kommt, gruppieren sich die anderen Varianten der Durchführung um einen Mittelwert von 4,5 (siehe Tabelle 3 und Abb. 2). Dementsprechend fällt die Einfaktorielle Varianzanalyse (ANOVA) hoch signifikant aus (Sig. $<.001$) und bestätigt das grafisch sichtbare Ergebnis. Paarweise Mittelwertvergleiche (Post-hoc-Tests: Bonferroni und Student-Newman-Keuls) zeigen, dass es die Variante Onlinelehre mit haptischem Planspiel ist, die sich signifikant von allen anderen Varianten unterscheidet, während sich die anderen drei Varianten nicht voneinander unterscheiden. Onlineveranstaltungen mit haptischen Planspielen werden von Studierenden merklich schlechter bewertet als die anderen Durchführungsvarianten. Ausschlaggebend ist die offensichtlich inadäquate Zusammensetzung aus digitaler Lehre und eigentlich haptischem Spielmaterial. Trotz aufwändiger Transformation der Brettspiele in den digitalen Raum, können "haptische" Spiele ihr Potential bei Onlineveranstaltungen nur eingeschränkt entfalten, was sich in der Zufriedenheit bemerkbar macht. So fällt für die Gesamtbewertung (Lernen und Zufriedenheit) bei haptischen Spielen neben dem größten in dieser Studie berichteten Mittelwertsunterschied (0.52 Punkte) auch die größte Effektstärke auf. Mit $d=.474$ zeigt sich ein annähernd mittlerer Effekt, der sich von allen anderen Effektstärken abhebt (Tabelle 3). Demgegenüber unterscheiden sich Lehrveranstaltungen mit computerbasierten Spielen nicht signifikant hinsichtlich Präsenz- oder Onlinedurchführung (Tabelle 3).

Abb. 2 Gesamtbewertung und Lernen nach Art der Planspieldurchführung



Diskussion und Fazit

Mit Beginn der pandemiebedingten Onlinelehre mussten innerhalb kurzer Zeit Konzepte für die Durchführung von planspielbasierten Lehrveranstaltungen entwickelt werden. Anhand von Evaluationsdaten möchte dieser Beitrag aufzeigen und diskutieren, inwiefern das gelungen ist, und wo sich Gemeinsamkeiten und Unterschiede zwischen Präsenz- und Onlinelehre zeigen.

Planspiele sind eine komplexe Lehr-Lernmethode, in deren Umsetzung verschiedene Komponenten zusammenkommen müssen: technische Voraussetzungen, Raum für Teamarbeit, inhaltliche und technische Einführung in die Simulation, Möglichkeit zur Interaktion zwischen einzelnen Teams und Lehrenden sowie Möglichkeit zur Interaktion aller Teams mit Lehrenden im Plenum. Trotz dieser hohen Hürden zeigen die Daten, nur geringe Unterschiede zwischen synchroner Präsenz- und Onlinelehre. Auf den ersten Blick mag dieses Ergebnis erstaunen und so stellt sich die Frage nach den Gründen. In einem empirischen Vergleich von Fernstudienangeboten mit Präsenzangeboten, folgern Gerhard et al. (2015), dass nicht die Unterscheidung von Präsenz/Online ausschlaggebend sei, sondern die didaktische Aufbereitung des Lernmaterials und

die Qualität der medialen Umsetzung. Lehner und Sohm (2021) greifen diesen Gedanken auf und betonen, dass vor allem die “qualitative Ausgestaltung” der Lehre wichtig sei und dies unabhängig von der Frage, ob in Präsenz oder Online. In einer qualitativen Forschungsarbeit zu Merkmalen guter Onlinelehre aus studentischer Perspektive wird deutlich, dass insbesondere die soziale Interaktion mit Lehrenden (Zeit nehmen, Erreichbarkeit, auf Rückfragen eingehen) als ausschlaggebend angesehen wird (Hawlitschek et al., 2022). Insbesondere in synchroner Onlinelehre kann soziale Interaktion mit Videokonferenzsystemen (Echtzeitübertragung von Ton und Bild) gut gelingen. Auf Knopfdruck können alle Beteiligten zwischen Plenum und Kleingruppen wechseln und so in unterschiedlichen Zusammensetzungen kommunizieren. Kirchner stellt gar die Frage, ob bei Onlinelehre mit Echtzeitübertragung von Ton und Bild nicht von Präsenz gesprochen werden könne (Kirchner, 2021, S. 261). Die Unterscheidung von synchroner Präsenz- und Onlinelehre ist demnach nur vordergründig relevant. In beiden Varianten kommt es auf die jeweils konkrete Ausgestaltung der Lehrveranstaltung und die Interaktion an. In der planspielspezifischen Literatur (Kriz, 2010; Roungas et al., 2018; Schwägele et al., 2021) wird häufig auf die Bedeutung von professionellem Facilitation für die erfolgreiche Durchführung von Planspielen hingewiesen. Auch dieser Aspekt ist nicht abhängig von Präsenz- oder Onlinelehre. In beiden Welten kann Facilitation anhand bestimmter Methoden und Kommunikationsformen gelingen. Die gleichzeitige physische Anwesenheit an einem Ort macht noch keine gute Lehre. Es sind die Lehrenden, die Art der Interaktion, die didaktische Aufbereitung und die vertiefte Auseinandersetzung mit einem Gegenstand, die gute Lehre ausmachen – in Präsenz und Online. Auch häufig diskutierte Hürden der Onlinelehre, wie soziale Kontakte oder hohe Anforderungen an eigenständiges Lernen (vgl. Kap. 3) sind in der synchronen Onlinelehre weniger bedeutsam als in asynchronen Veranstaltungen. Die Echtzeitübertragung von Ton und Bild ermöglicht eine hohe soziale Einbindung, sowohl zu Peers (siehe Skalen zur Teamarbeit, Forschungsfrage 1) als auch zu Lehrenden (siehe Skala Dozierende: Kommunikation, Forschungsfrage 2). Evtl. ist die Unterscheidung von synchron und asynchron für Studierende wesentlich einschneidender als die Unterscheidung von Präsenz und Online. Denn gerade in asynchronen Lernsettings ist Eigenständigkeit gefragt, bei gleichzeitiger Abwesenheit von Lehrenden und Peers.

Als Fazit kann festgehalten werden, dass synchrone Onlinelehre mit Planspielen gut gelingt. Teilweise finden sich für die Onlinelehre sogar leicht bessere Ergebnisse als für die Präsenzlehre: Zum Beispiel wird die Kommunikationsfähigkeit der Dozierenden besser eingeschätzt. Auch wenn Studierende bei Betrachtung aller Planspielveranstaltungen Präsenzlehre hinsichtlich Lernen und Zufriedenheit etwas mehr schätzen als die Onlinelehre, so zeigen detailliertere Analysen, dass dieser Effekt auf die Konstellation Onlinelehre mit haptischen Planspielen zurückzuführen ist. Insgesamt bewegen sich die gefundenen Unterschiede auf einem eher niedrigen Niveau (nur kleine Effektstärken).

Die Unterscheidung von Präsenz und Online ist für die Frage nach "guter Lehre" demnach wenig relevant. Ob Lehre in Präsenz oder Online gehalten wird, ist aber sehr relevant für die Hochschulentwicklung und Organisation von Lehre in der Zukunft, etwa im Hinblick auf Weiterbildungsangebote für Lehrende und die materielle (und/oder digitale) Ausstattung von Hochschulen. Die Erfahrungen aus der pandemiebedingten Onlinelehre sollten genutzt werden. Insbesondere aus der Ortsunabhängigkeit der Onlinelehre ergeben sich strategische Möglichkeiten. So kann Hochschul- und Studiengangsentwicklung in fakultäts-, hochschul-, und sogar länderübergreifenden Kooperationen gedacht werden.

Restriktionen

Auch wenn die große Datenbasis empirisch gut abgesicherte Schlussfolgerungen erlaubt, sind doch einige Restriktionen zu beachten.

Die Daten wurden ausschließlich in dualen Studiengängen an der Fakultät Wirtschaft und Gesundheit der DHBW Stuttgart erhoben. Die Ergebnisse können daher nicht zwingend auf andere Hochschularten und Fachbereiche übertragen werden. Hinzu kommt ein institutioneller Aspekt: Am ZMS ist Wissen um planspielbasierte Lehre institutionell gebündelt und wird regelmäßig in Workshops für Planspielleitungen vertieft und geteilt. Zu Beginn der Pandemie wurden zügig didaktische Konzepte entwickelt und erprobt. Zudem ist ein Großteil der Lehrenden am ZMS seit Jahren in der planspielbasierten Lehre aktiv. Auch diese institutionelle Komponente mag dazu beitragen, dass Planspiele in der Onlinelehre ähnlich gut bewertet werden, wie Planspiele in der Präsenzlehre. Die vorliegende Auswertung

basiert zudem auf der Selbsteinschätzung von Studierenden. Auch wenn Studien die hohe Urteilskompetenz von Studierenden belegen (vgl. Cronan et al., 2012; Rindermann, 2003), so geht dies doch mit Problemen einher (Spinath & Seifried, 2018), die bei der Bewertung der Ergebnisse berücksichtigt werden sollten.

Literaturverzeichnis

- Broadbent, J. & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13.
<https://doi.org/10.1016/j.iheduc.2015.04.007>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Lawrence Erlbaum Associates.
- Cronan, T. P., Leger, P.-M., Robert, J., Babin, G. & Charland, P. (2012). Comparing Objective Measures and Perceptions of Cognitive Learning in an ERP Simulation Game: A Research Note. *Simulation & Gaming*, 43(4), 461–480.
- Entner, C., Fleischmann, A. & Strasser, A. (2021). Hochschullehre im digitalen Wandel. Überlegungen zur didaktischen Gestaltung von Präsenz- und Onlinelehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba, M. Wiemer & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre*. DUZ Verlags- und Medienhaus GmbH.
- Fallmann, I., Steinbacher, H.-P., Ammenwerth, E., Burian, R., Ebner, M., Ghoneim-Rosenauer, A., Gegenburger, B., Grün, E., Hauser, W., Heinrich, M., Karpf, K., Krizek, G., Lietze, S., Mantler, H., Mair, M., Niederl, F., Oberhumer, P., Herwig, R., Rinnhofer, C., . . . Zitek, A. (2021). *Quantifizierung von virtueller Lehre an österreichischen Hochschulen*. Verein Forum neue Medien in der Lehre Austria.
- Geithner, S. & Menzel, D. (2016). Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. *Simulation & Gaming* 47(2), 228–256. <http://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-353942>
- Gerhard, D., Heidkamp, P., Spinner, A., Sommer, B., Sprick, A., Simonsmeier, B. A. & Schneider, M. (2015). Vorlesung. In M. Schneider & M. Mustafić (Hrsg.), *Gute Hochschullehre: Eine evidenzbasierte Orientierungshilfe* (S. 13–35). Springer Berlin Heidelberg.
- Hawlitsek, A., Briese May, S. & Albrecht, P.-G. (2022). „Man fühlt sich nicht alleine gelassen.“ Merkmale guter Online-Lehre aus studentischer Perspektive. *die hochschullehre*, 8(3), 31–44.
- Kirchner, A. (2021). Unendliche Räume - ein Planspiel in digitaler Präsenz. *die hochschullehre*, 7(24), 251–264.
- Kriz, W. C. (2010). A Systemic-Constructivist Approach to the Facilitation and Debriefing of Simulations and Games. *Simulation & Gaming*, 41(5), 663–680.
<https://doi.org/10.1177/1046878108319867>

- Kriz, W. C. (2011). Qualitätskriterien von Planspielanwendungen. In S. Hitzler (Hrsg.), *ZMS-Schriftenreihe: Bd. 2. Planspiele - Qualität und Innovation: Neue Ansätze aus Theorie und Praxis* (1. Aufl., S. 11–38). Books on Demand.
- Lehner, M. & Sohm, K. (2021). Qualität, didaktische Methodik und Digitalität: Erfahrungen und Reflexionen an der Fachhochschule Technikum Wien. In U. Dittler & C. Kreidl (Hrsg.), *Wie Corona die Hochschullehre verändert* (S. 339–350). Springer Fachmedien Wiesbaden.
- Mandl, H. (2004). On the path to a new learning culture - the contribution of web-based business games. In T. Eberle (Hrsg.), *Bridging the gap: transforming knowledge into action through gaming and simulation: Proceedings of the 35th Conference of the International Simulation and Gaming Association, Munich, 2004* (S. 163–174).
- Rachbauer, T. & Hanke, U. (2022). Hybride, blended synchronous und Hyflex-Lehre – Chancen, Risiken und Gelingensbedingungen. *Zeitschrift für Hochschulentwicklung*, 17(2), 43–60. <https://doi.org/10.3217/zfhe-17-02/03>
- Reinmann, G. (2016). Gestaltung akademischer Lehre: semantische Klärungen und theoretische Impulse zwischen Problem- und Forschungsorientierung. *Zeitschrift für Hochschulentwicklung*, 11(5), 225–244.
- Reinmann, G. (2021). *Hybride Lehre - Ein Begriff und seine Zukunft für Forschung und Praxis*. Impact Free Journal für freie Bildungswissenschaftler. https://gabi-reinmann.de/wp-content/uploads/2021/01/Impact_Free_35.pdf
- Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussforderungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation*, 2(2), 233–256. <https://www.degeval.org/en/journal-of-evaluation/volumes/heft-22003/>
- Roungas, B., Wijse, M. de, Meijer, S. & Verbraeck, A. (2018). Pitfalls for Debriefing Games and Simulations: Theory and Practice. In A. Naweed, M. Wardaszko, E. Leigh & S. Meijer (Hrsg.), *Lecture Notes in Computer Science. Intersections in Simulation and Gaming* (Bd. 10711, S. 101–115). Springer International Publishing. https://doi.org/10.1007/978-3-319-78795-4_8
- Schwägele, S. (2015). *Planspiel - Lernen - Lerntransfer*. Dissertation. *ZMS-Schriftenreihe: Band 7* [348 Seiten]. Books on Demand GmbH.
- Schwägele, S., Zürn, B., Lukosch, H. K. & Freese, M. (2021). Design of an Impulse-Debriefing-Spiral for Simulation Game Facilitation. *Simulation & Gaming*, 52(3), 364–365. <https://doi.org/10.1177/10468781211006752>
- Simons, J., Leverett, S. & Beaumont, K. (2020). Success of distance learning graduates and the role of intrinsic motivation. *Open Learning: The Journal of Open, Distance and e-Learning*, 35(3), 277–293. <https://doi.org/10.1080/02680513.2019.1696183>
- Spinath, B. & Seifried, E. (2018). Was brauchen wir, um solide empirische Erkenntnisse über gute Hochschullehre zu erhalten? *Zeitschrift für Hochschulentwicklung*, 13(1), 153–169. <https://zfhe.at/index.php/zfhe/issue/view/57>
- Torun, E. D. (2021). Online Distance Learning in Higher Education: E-Learning Readiness as a Predictor of Academic Achievement. *Open Praxis*, 12(2), 191. <https://doi.org/10.5944/openpraxis.12.2.1092>
- Trautwein, F. & Alf, T. (2022). Theoriebasierte Entwicklung eines Inventars zur Evaluation von Planspielveranstaltungen. In T. Alf, S. Hahn, B. Zürn & F.

Trautwein (Hrsg.), *ZMS-Schriftenreihe. Planspiele - Erkenntnisse aus Praxis und Forschung: Rückblick auf den Deutschen Planspielpreis und das Europäische Planspielforum 2021*. Books on Demand.

Xu, D. & Xu, Y. (2020). The Ambivalence About Distance Learning in Higher Education. In L. W. Perna (Hrsg.), *Higher Education: Handbook of Theory and Research. Higher Education: Handbook of Theory and Research* (Bd. 35, S. 351–401). Springer International Publishing. https://doi.org/10.1007/978-3-030-31365-4_10

5.3 The Role of Reflection in Learning with Simulation Games – A Multi-Method Quasi Experimental Research

Abstract

Introduction: The role of reflection in experience-based learning is discussed widely. This paper researches the effects of a reflection assignment as part of debriefing with qualitative and quantitative methodology in five simulation game-based seminars.

Research Design and Methods: The intervention (reflection assignment) used in this study consists of three reflective questions that are discussed in groups. We aim to find out the effects of adding this intervention to the regular in-between-debriefings (three times during the whole seminar). The methodological setup is quasi-experimental and so, consisted of a test group and a control group. Two different types of simulation games were used, one a general management game and the other a change management game. Data used in this study is gathered from three different sources: First, we used in-game-performance data from both games. Second, we used two types of questionnaires to evaluate game experience and self-reported learning (MEEGA+ and ZMS inventory). Third, the reflection notes from the treatment-teams were used for qualitative analysis.

Results: In several dimensions (game-success, game-experience and self-reported learning) significant differences were found between treatment and non-treatment students. Qualitative data show a deep level of reflection for treatment teams differentiating for the two simulation games evaluated.

Discussion: Supported by the qualitative reflection data we assume that the better results for treatment-students are routed in the repetitive reflection. The permanent circular reflection helps treatment students to understand the games better and gain deeper insights.

DOI: <https://doi.org/10.1177/10468781231194896>

Autoren: Tobias Alf, Marieke de Wijse, Friedrich Trautwein

Introduction

In this research the effects of a formative assessment method consisting of three reflective questions is evaluated using a mixed method research design. In five sg based seminars treatment groups received an additional reflection assignment while non-treatment groups had “only” the normal debriefing. Differences between treatment and non-treatment groups are evaluated in various ways and qualitative analysis shows how reflection adds value for learning. The sgs used in this study are TOPSIM General Management (GM) and Riva Sys Teams Change (STC). To introduce the theory behind our research we draw the connection between reflection and learning and describe why this connection is relevant for learning with sg. After introducing the theory on levels of learning we outline the research-design and detailed research questions (RQ) for this study. So then we describe in detail how each RQ is analyzed (methods). After presenting the results of this study per RQ we discuss our findings and name limitations.

Introduction to the role of reflection in learning

The idea of sgs as an experiential learning method is “to provide an enhanced learning experience” (Pasin & Giroux, 2011, p. 1243). A simulated learning environment gives students the chance to experience, reflect and learn. Usually, sgs follow the structure of briefing, experience and debriefing (Duke, 2014; Geurts, Duke, & Vermeulen, 2007; Klabbers, 2009; Stoppelenburg, Caluwé, & Geurts, 2012). Briefing serves to introduce students to the simulation and prepare them with what they need to know about the sg (scenario, rules, roles and resources etc.) to start playing it. Part of briefing can be establishing goals for learning which can contribute to reflection later on. While the sg is played students act in their roles and gather experience, which then serves as input for reflection and debriefing both during and after the sg (Schwägele, Zürn, Lukosch, & Freese, 2021).

In literature sgs are associated with experience- and action based learning theories (Kolb & Kolb, 2009; Taylor, Backlund, & Niklasson, 2012), for example Kolb’s Experiential Learning Cycle (Alklind Taylor, 2014; Geithner & Menzel, 2016). It describes learning as a continuing cycle of concrete experience, reflection, abstract conceptualization and active experimentation (Kolb & Kolb, 2009).

Learning can occur when experience is reflected and conceptualized. The new concepts are transferred to action, which leads to new experiences. Argyris and Schön developed another systemic oriented experiential learning theory and identified the concept of first and second order learning (1974) which is described in detail below. Schön adds the concept of reflection-in-action and reflection-on-action (1983). He elaborates how professionals handle (challenging) situations based on their experience by immediate reflection during their action (reflection-in-action) and reflecting on it afterwards (reflection-on-action). Relating to the concept of Schön, Greenwood adds the idea of reflection-before-action, meaning to anticipate a situation beforehand (1993).

Many authors emphasize the relationship between experience and reflection in experiential learning theories (Alklind Taylor, 2014; Hilzensauer, 2008; Klabbers, 2009; Kolb & Kolb, 2009; Nakamura, 2022). In sum, sgs offer opportunities to learn from experience through reflection. In the game play phases students aim to solve simulated challenges and thereby gather experience as an input for reflection/debriefing.

Contemporary research on reflection in simulation games

Reflection has a role in the learning processes within sgs (Baker, Jensen, & Kolb, 1997; J. Y. Lee, Donkers, Jarodzka, Sellenraad, & van Merriënboer, 2020; Nakamura, 2022). Adding reflection assignments during the gameplay is a form of formative assessment which enables participants to use the learnings from reflection and utilize the sg to the full. Only reflecting in debriefing after gameplay would leave out many opportunities for learning, whereas adding reflection triggers and opportunities during a sg session can serve as leverage points for learning. Specific research on the different forms of reflection and their effects in sgs is scarce at the moment. Nakamura (2022) researched the impact of adding pre-structured questions before the gameplay and in reflections and found a positive significant result on learning. Lee et al. (2020) performed a medical simulation study in which participants could take time outs resulting in higher experienced agency but not in an increased learning outcome. The authors explain the result by concluding participants did not receive specific instruction or prepared questions on how to use this reflection time effectively. Husebø (2013) researched the effects of debriefing questions on reflection during debriefing after gameplay with

participants in medical sgs and found a positive (qualitative) result and suggests further research is necessary on types of questions asked and their effects on learning. We do not know if these results are generalizable to sgs in general, since medical sgs often have specific contexts such as procedural orientation with given norms on what behavior is correct according to pre specified procedures. However, there are also medical sgs having more open characteristics for instance when aimed at more tacit knowledge and competencies such as situational awareness.

These findings indicate structured questions could add value to (in-between-) debriefing to increase learning, however the number of studies so far is limited. This study aims to contribute to researching the effects of structured reflection assignments on learning with sgs.

Theoretical background on levels of learning

Argyris and Schön identified the concept of first and second order learning (1974) which has been used and referenced widely in management literature (Bartunek, 2014; Crossan, 2003; Friedman, Torbert, Nielsen, Silverman, & Bradbury, 2014). The basis of systemic oriented learning is seen in reflecting on and responding to deviations that happen in a given system. The type of response determines the type of learning involved. Theory development from the founding father of systems learning Chris Argyris is applicable here. First order learning can be defined as learning to apply predetermined rules and procedures in a given system. It leads to successful actions within the system. Second order learning requires a reflection on the norms behind the decisions and assumptions so processes can be adapted in such a way they fit the requirements of the situation. These process interventions are directed at checking if the norms applied are still relevant and adequate for the challenges the people in the organization are faced with (Argyris, 1977; Greenwood, 1988; Tosey, Visser, & Saunders, 2012). Greenwood (1988) emphasizes that especially second order learning is based on reflection because it criticizes the underlying structures and norms behind them. The concept of third order learning was introduced based on the work of Argyris as an extension of his theory (Tosey et al., 2012; Visser, 2007; Visser, Chiva, & Tosey, 2018). Third order learning means to reflect on how actors take on their roles in the system and if this is adding value to their personal values/needs and the values/needs of the system. In addition, third order learning is about learning to learn and remain adaptive. In

this process also less tangible issues regarding virtues of wisdom and justice can be involved (Reynolds, 2014). Applying these levels of learning to the qualitative part of this study (especially research question 3 see next paragraph) enables us to draw conclusions on what type of learning was most prevalent and if this was in line with the aim of the specific case.

Aim of research and research questions

The aim of this research is to evaluate the impact of a structured reflection assignment that is repeated three times during the sg-based seminars as a part of the roundly in-between-debriefings. Therefore, treatment groups participated in a reflection assignment while non-treatment groups did not. The groups are then compared in various aspects using quantitative and qualitative methods. In detail the following research questions (RQ) are addressed in this study:

- RQ1: How does the reflection intervention affect the sg results and perceived learning in the treatment groups?
- RQ2: How does the reflection intervention affect game play variables (assessment of in-between-debriefing and debriefing, student engagement, teamwork, assessment of simulation)?
- RQ3: What levels of learning/reflection do we see in the notes on the flips and can there be made interpretations on in what proportion they occur?
- RQ4: What learning process do we see in the teams from round to round?

Research design and methods

Overview research design

To address the formulated RQ a mixed method setup was used (see figure 1). Quantitative evaluation was applied with two post game questionnaires and qualitative analysis was applied to the reflection notes gathered. Five sg-based seminars were evaluated with a quasi-experimental design. Altogether 110 students participated in 25 teams. In each seminar two groups received a special reflection assignment with pre structured reflection questions (treatment groups) that was repeated three times during gameplay (beginning, middle, end). The other two to four groups (depending on how many students attended the seminar) did not receive the reflection assignment. In the treatment groups the assignment

became part of the in-between-debriefing and was added after the discussion of results in the plenary. For non-treatment groups the in-between-debriefing ended after the plenary discussion. So, these groups had a little (about 10 to 15 minutes) more time to read, calculate and discuss their next round's decisions within their teams. The treatment was implemented three times during the sg-based seminars: At the start (after round one), in the middle of gameplay (round 3 for GM and round 2 for STC) and towards the end of the game (round 5 for GM, round 3 for STC). The procedure differentiated slightly because GM consists of six rounds and STC of three. There was no interaction between the groups at the reflection moments. The reflection-assignment exists of three pre structured questions that guide the teams to reflection on their decisions and behavior during the sg:

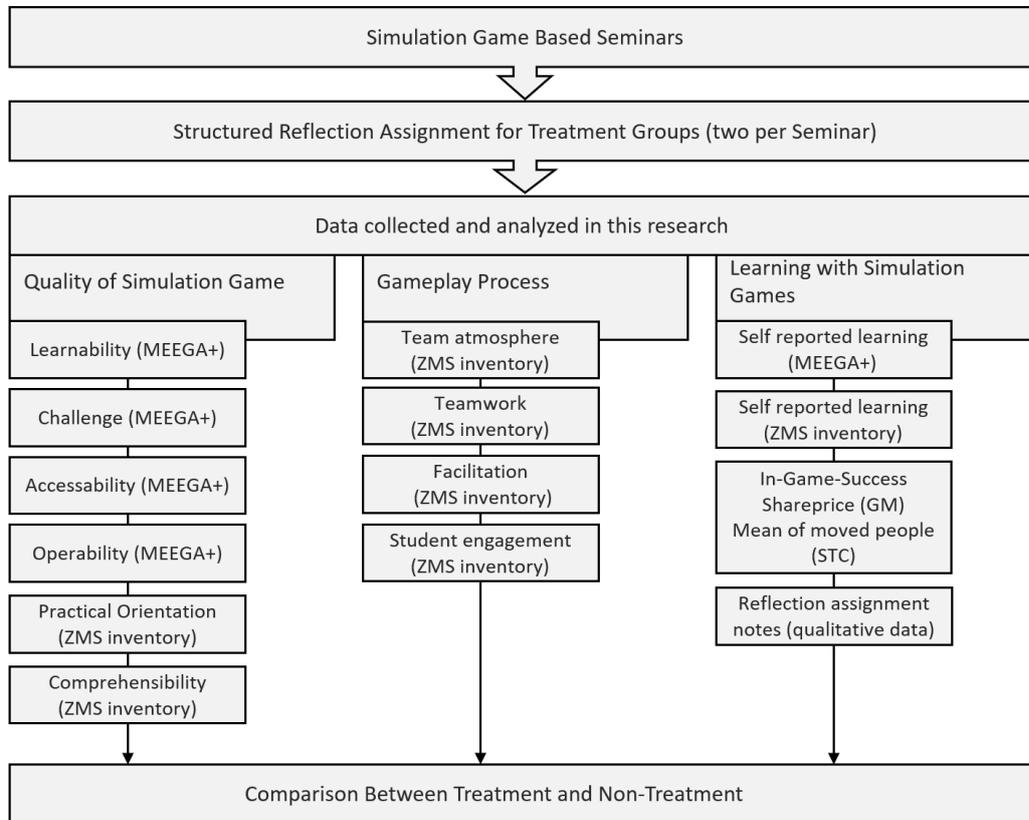
1. What went well?
2. What went not so well?
3. What do I need, what do we need to reach our goals?

These pre structured questions were open and quite general to allow participants the opportunity for personalized input on what they found relevant. Previous studies (J. Y. Lee et al., 2020; Nakamura, 2022) indicate that pre-structured questions support learning via reflection and also preparation with goal setting for the next game round.

In onsite seminars the teams received a prepared flipchart with invitation to discuss the questions written on the flipcharts and write their answers on the chart. In other case studies (Wijse-van Heeswijk, Rouwette, & van Laere, in press) we noticed participants were not so inclined to fill out the questions if they were not pre written on the flip over. In online seminars the teams received a prepared PowerPoint sheet with the same questions and were invited to discuss answers and make notes by screen-sharing. Afterwards the results were presented briefly to the professor and the other team.

At the start of the seminars students were informed that the seminar was evaluated in the context of a research project on the role of learning in sgs. Participation in the study was voluntary and data was gathered anonymously using Evasys Version 8.2.

Figure 1: Overview Research Design



Quantitative analysis approach (RQ1 and RQ2)

RQ1 focuses on the output of the seminar (learning and game results) while RQ2 focuses on the gameplay process (teamwork, engagement, assessment of simulation). To address these RQs students were asked to answer a longer questionnaire at the end of the seminars. These questions were aimed at reflecting on the seminar from start to end. The questionnaire consists of two inventories that are described below. Both inventories consider self-reported learning results (RQ1) as well as gameplay-process variables that are addressed in RQ2 (see figure 1 for an overview). The first part of the questionnaire is called the ZMS inventory and was developed at the Center for Management Simulation. It measures learning satisfaction and relevant influencing factors on learning with simulation games (Trautwein & Alf, 2022) and consists of 27 Likert scaled items that can be combined to seven scales. Two scales measure the quality of the simulation game regarding comprehension of the game and the relation to realism of the game. One scale of

five items measures how the simulation is facilitated. Two scales measure different aspects of teamwork (task orientation and atmosphere). Another scale measures the student engagement during the seminar and the final scale is about overall satisfaction and learning. Exploratory and confirmatory factor analysis confirmed the seven scales used in this research and alpha values between .783 and .943 can be reported for the scales (Trautwein & Alf, 2022). The second questionnaire is the MEEGA+ instrument developed by Petri et al. (2017) to evaluate educational games. The MEEGA+ questionnaire is a validated instrument and was originally developed for games with digital interfaces. Therefore, some questions were adapted to the games used in this study. For example, items used to evaluate the accessibility of games ask about fonts and colors. In the context of this study we asked “The game interface is clearly arranged.” and “The user interface is simple and well structured.” The social interaction scale of MEEGA+ asks if there was the possibility of interaction which is natural for the sg used in this research. So, we converted the question to whether the simulation promotes cooperation with other students. For quantitative analysis treatment students and non-treatment students are compared (for some analysis treatment groups and non-treatment groups are compared using group means). Since evaluation data on group level is $n < 30$ Mann-Whitney-U-tests are calculated for test statistics. U-Tests are appropriate for smaller samples which we have on group level and are even appropriate for non-parametric data (Bortz & Schuster, 2010; Eid, Gollwitzer, & Schmitt, 2011, p. 322). To address RQ1 also success indicators from the sg are used knowing that this is controversially discussed (Brazhkin & Zimmerman, 2019; Pasin & Giroux, 2011). Critique on the use of game success indicators is that in some games groups could perform well by random luck instead of learning. We find this is very unlikely for the sgs used in this study because the simulations used have a high level of complexity. For GM students have to make many interdependent decisions in six consecutive rounds¹³ which means that they have to bring their decisions into a coherent overall concept. Also, there is not one right solution but multiple strategies possible which have to be performed in a coherent way to achieve game success.

¹³ E.g. if it is decided to lower the price for a product students can assume to sell more of the product. This means that they have to produce more. Producing more means to hire new workers and to buy new machines. New machines have to be financed. Therefore the decisions are interrelated.

For STC students have to play 42 action cards in a logical order according to the simulated change management process. For many action cards it is required to select proper game characters to play them which has a certain level of difficulty. Since many different characters in the game have to be approached it is highly unlikely one would select the right candidates by luck for each intervention every time. In both simulations it can be assumed as very unlikely that successful decisions can be taken by chance permanently. If a team performs well in the game, it indicates at least that they have learned to a certain extent how to play the sg in line with the learning goals of the sg.

Since RQ2 addresses process variables such as the quality of in-between-debriefings a short questionnaire was conducted after each in-between-debriefing. Students were asked to answer the perceived quality of debriefing using 3 items that were combined to a scale. Treatment groups filled it out after their assignment, control groups after the joint debriefing discussion. Via this procedure it is possible to compare the treatment groups with the non-treatment groups on the perceived quality of in-between-debriefing.

Qualitative analysis approach (RQ3 and RQ4)

In addition to quantitative data from questionnaires and game performance data the research design also provides qualitative data from the treatment groups. The notes from the reflection assignments were collected, documented, and used for qualitative analysis. This may grant insight into the level of group reflection and by comparing consecutive reflections of groups in learning processes. With RQ3 we look at the data on a micro level analyzing the single answers given by groups. With RQ4 we analyze the data in a broader view comparing the notes in a chronological manner.

With regard to RQ3 we converted the text from the flaps to an excel file and assigned codes according to the coding scheme in the appendix (Argyris, 1977; Visser et al., 2018): Code 1 for first order learning, Code 2 for second order learning and Code 3 for third order learning. Using these codes, we analyze the reflective response to find out what teams were involved with what learning levels during the seminar (these could either be actual learnings or reflections on certain learning levels). As described above, treatment teams participated in three reflection interventions during the process of gameplay answering three open

questions (after round one, in the middle of the game and towards the end of the game).

Comparing the discussion-notes (flipcharts or digital notes) in a chronological manner gives insight into a process of discussion and learning the groups went through (RQ4). Applying the idea of process tracing (Cloutier & Langley, 2020; Collier, 2011; Langley, 1999; Langley & Tsoukas, 2016; Munro, 2016) we analyzed the data to find relevant chronological patterns. Therefore, five relevant scenarios of possible processes were formulated, marked with colors and the occurrence of scenarios was counted as is shown in the following table.

Table 1: Description of the five possible learning scenarios

	Short Deskription	Description
Scenario 1	Topic occurs only in one period	Issue was identified in one period, but unclear how it was processed. This doesn't necessarily mean participants did not learn how to deal with it but from the data we have we cannot deduct what happened to it.
Scenario 2	Positive → positive	Issue was identified as "worked well" in earlier rounds. Team managed the issue continuously in a satisfying way.
Scenario 3	negative → positive	Issue was identified as "worked not so well" or "What do we need to...". Team managed to work on the issue so that it was identified as positive in later rounds.
Scenario 4	negative → negative	Issue was identified as "worked not so well" or "what do we need to..." in earlier rounds. Team was not able to work on the topic in a satisfying way so that it was identified as negative again.
Scenario 5	positive → negative	Issue was identified as "worked well" in earlier rounds. Team had problems dealing with the topic in later rounds.

Using the described scenarios, we see which topics were discussed when in the seminar and how the topics proceeded during the seminar. For example, scenario 3 negative → positive shows that teams identified an issue as relevant but have so far not managed to work on the issue in a satisfying way ("went not so well" or "what do we need..."). In later rounds the topic is identified as "worked well", indicating that teams found solutions

to manage the issue properly. Especially scenario 3 can be interpreted as an indication for learning. By working on the issue it turned from negative to positive, indicating that students learned how to deal with the problem¹⁴.

Case study and context description

Using the described methodology, we evaluated five simulation based seminars at the DHBW Stuttgart Business Faculty that were taught at the Center for Management Simulation (ZMS). The Center for Management Simulation is a specialized lab for sgs. The setup of the management simulation is as follows: In a plenary room the briefing, in-between-debriefing and final debriefing took place, during gameplay the teams have their own offices with a big screen to share data, whiteboards and flaps to share notes and organization. Due to the corona pandemic only two of the evaluated seminars were conducted on site (see Table 2). Three other seminars were conducted via MS Teams where teams had their virtual rooms. We evaluated four seminars using the simulation game TOPSIM General Management which is a complex business sg that has been used in scientific study before (Hühn & Rausch, 2022). Students take on the role of the board of a company and make decisions in six areas of management (TOPSIM GmbH). All teams start at the same position and compete with each other on one market. The most general performance indicator is the share price of a company. All four seminars were facilitated by an experienced professor using the simulation for approximately 18 years. In the fifth seminar a simulation on change management (riva solutions GmbH) was used. Students take over the role of external change consultants and guide a middle size organization through a change management process. The goal of the game is to move a representation of 26 characters from total rejection of the change-process to acceptance of the process. The simulation was facilitated by two experienced trainers that taught this

¹⁴ Though we are not excluding that the other scenarios may activate learnings with stimuli in the long run maybe even after the sg by recognizing how the negative feedback can aid them in drawing positive learnings.

specific simulation more than ten times. The actual plan was to evaluate at least two change management seminars. Due to the corona semester this was not possible. But it provides at least one example of another sg to discuss whether the effects are typical for GM or can also be generalized to other sgs. Both sg's are round based simulations. After each round the simulations were "interrupted" for an in-between-debriefing discussion in plenary, that is led by the lecturers. Relevant topics and results of the last round were discussed and students had the opportunity to ask questions. After the in-between-debriefing in the plenary the treatment teams started their reflection assignment.

Table 3: Overview of seminars and sg used for evaluation

	Online/on site	Simulated Rounds	Simulation	Teams	Reflection assignment
Seminar 1	on site	6 of 6	TOPSIM General Management	4 total (2 treatment)	after round 1, 3 and 5
Seminar 2	online	6 of 6	TOPSIM General Management	5 total (2 treatment)	after round 1, 3 and 5
Seminar 3	online	6 of 6	TOPSIM General Management	5 total (2 treatment)	after round 1, 3 and 5
Seminar 4	online	6 of 6	TOPSIM General Management	5 total (2 treatment)	after round 1, 3 and 5
Seminar 5	on site	3 of 3	Riva SysTeamsChange	6 total (2 treatment)	after round 1, 2, 3

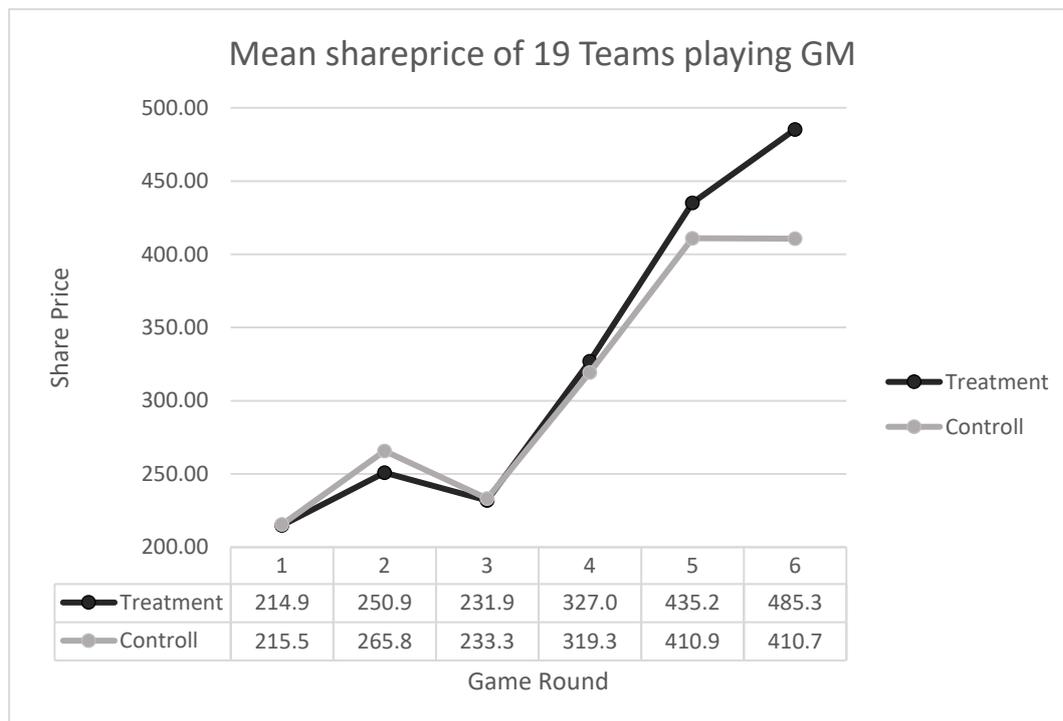
Results/findings from both games

In this chapter the results of our research are described. In the text we follow the structure of the RQ raised above.

RQ1: Effects on performance in the game and perceived learning

Comparing treatment teams and control teams for GM (n=19) we see basically no effects in the first rounds. Starting from round 5 we see treatment-teams having better game results (Graph 1). Results in round six show a mean share price of 410 for non-treatment teams and a mean share price of 485 for treatment teams¹⁵.

Figure 2: game performance of treatment and control teams for GM



The game success measures for STC (n=6 teams) indicate that the two treatment-groups have the best results and the lowest standard deviation (Table 3). For STC the standard deviation is an important measure because it shows that a team was able to move all 26 characters simultaneously by their change management activities.

¹⁵ T-Test result is slightly not significant (.056), which may be due to the small sample of team means (n=19).

Table 3: Game Success measures SysTeamsChange

	Treat- ment	Treat- ment	Control	Control	Control	Control
Success (average of moved people on the game board)	21,27	21,42	16,62	18,46	20,39	19,5
Standard deviation	1,82	1,20	4,45	3,66	2,40	1,20

Looking at the results of both sgs, we see the tendency of better results for the treatment groups. Using z-standardization (Pospeschill, 2022, p. 190) the success measures of both sgs were brought together to one scale. Due to the small sample of team means (n=25) a Mann-Whitney-U-Test for independent samples was calculated to identify significant differences between reflection groups and non-reflection groups. The test shows a highly significant result (.008) and a medium to strong effect size of $r=.521$ (Table 4). It confirms what is seen looking at both games' individual success measures. There is a tendency of treatment groups having better results.

Interpreting students' perceived learning variables, a highly significant result is found on the MEEGA+ learning scale but not on the ZMS learning scale (Table 4). In detail three out of the four MEEGA+ learning single items show significant results: Treatment students assessed their learning better regarding cost calculation, operations management and for interrelated business knowledge (Table 4). The conflicting results between learning on MEEGA+ and the ZMS inventory may be due to the fact that the ZMS scale is formulated very general and consists of 6 items measuring learning and satisfaction on a general level (e.g. "I have learned a lot in the simulation game" or "How satisfied are you with the course overall?"), whereas the variables of MEEGA+ are formulated very specific¹⁶ and request students to assess their learning with regard to a very specific topic. From the results can be deduced that treatment groups and non-treatment groups

¹⁶ The researcher using MEEGA+ is asked to fulfill the sentence "The game contributed to..." with regard to specific learning goals (Petri et al 2018, Springer)

assessed the sg not different concerning learning and satisfaction (ZMS inventory)
but made a difference when they were asked for detailed learning goals.

Table 4: Results on game success and learning (RQ1)

RQ1

Scale or Item		N	Mean Rank	Sum of Ranks	Mann-Whitney-U-Test	Sig. (2-tailed)	Effect size r
z-standardized success measures GM and STC	Control	15	9,87	148,00	28,000	0,008	0,521
	Treatment	10	17,70	177,00			
	Total	25					
Learning (MEEGA+)	Control	65	48,14	3129,00	984	0,004	0,278
	Treatment	45	66,13	2976,00			
	Total	110					
Learning (ZMS)	Control	65	51,55	3350,50	1205,5	0,117	0,149
	Treatment	45	61,21	2754,50			
	Total	110					
Learning: Interrelated business knowledge (single item MEEGA+)	Control	49	36,66	1796,50	571,5	0,002	0,334
	Treatment	37	52,55	1944,50			
	Total	86					
Learning: Operations management (single item MEEGA+)	Control	49	38,32	1877,50	652,5	0,021	0,249
	Treatment	37	50,36	1863,50			
	Total	86					
Learning: Cost calculation (single item MEEGA+).	Control	49	38,70	1896,50	671,5	0,033	0,230
	Treatment	37	49,85	1844,50			
	Total	86					
Learning: Marketing (single item MEEGA+)	Control	49	39,45	1933,00	708	0,075	0,192
	Treatment	37	48,86	1808,00			
	Total	86					

RQ2: Effects on process and team variables

Comparing student's assessment on round based in-between-debriefings we do not see a significant result when comparing treatment students with non-treatment students. We can state that the reflection assignment did not improve students' perceived quality of the in-between-debriefings (Table 5). But we see a trend of treatment students to rate the debriefing better from round to round compared to non-treatment students (Table 5). While non-treatment students (control) rated the first in-between-debriefing slightly better, treatment students rated the second and third in-between-debriefing better with a tendency of greater differences towards the third measurement (end). This is also seen in the test-significance (becoming lower from time to time) and the effect size (becoming higher from time to time). Possibly the reflection assignment shows its effects only by repetition as we have also seen in RQ1. This interpretation must be made with caution because it might be due to coincidence as none of the tests show significant results.

Table 5: Assessment of debriefing with and without reflection assignment (RQ2)

In-between-debriefing		N	Mean Rank	Sum of Ranks	Mann-Whitney-U-Test	Sig. (2-tailed)	Effect size r
Start	Control	51	44,99	2294,50	918,5	0,830	0,023
	Treatment	37	43,82	1621,50			
	Total	88					
Middle	Control	40	33,41	1336,50	516,5	0,318	0,119
	Treatment	30	38,28	1148,50			
	Total	70					
End	Control	43	35,56	1529,00	583	0,253	0,132
	Treatment	32	41,28	1321,00			
	Total	75					

Comparing sg process variables from the ZMS inventory such as student engagement or teamwork we see significant effects between treatment and non-treatment students with small and medium effects (Table 6). Treatment students evaluate themselves to be more engaged (Sig. .024, $r=.216$) and to have both, a better team atmosphere (Sig. .000; $r=.351$) and a better task orientation (sig. .020, $r=.223$). So, differences between treatment and non-treatment students are not only seen in results as described for RQ1 but also in process variables that possibly explain the different results. From logic and several research review articles on effects of agency and motivation (Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012; Deen, 2015; Watt & Smith, 2021), we may infer that teams show better results if students are more engaged and perform better as a team. Sg comprehension is exact on the .05-significance-level with a rather small effect size, meaning that treatment-students rate themselves having a slightly better understanding of the sg than non-treatment students. This is consistent with other results. Having a better understanding of the simulation may lead to better results and learning. No significant results are found for the scoring on the lecturers of the seminars and the practical relevance of the simulation.

Looking at the MEEGA+ scales the results are in line with what is seen in the ZMS inventory. Among others we find significant differences for treatment and non-treatment students on the learnability scale and on the social interaction scale. Learnability here means to learn how to use and play the game (Petri, Gresse von Wangenheim, & Borgatto, 2018) and so has an intersection with the simulation variables from the ZMS inventory that ask for comprehension of the game. The structured reflection assignment likely leads to an improved understanding of the sg. In the treatment of our qualitative results we will go deeper into this explanation. The significant difference on the social interaction scale confirms the results from the ZMS inventory regarding teamwork. Both inventories suggest that the structured reflection assignments lead to an improved social experience as a group.

We conclude that the intervention did not influence the perceived quality of the roundly in-between-debriefing. But it has small to medium effects in three areas: treatment students are more engaged, they rate themselves to have a better understanding of the sg and an improved experience as a team. The three described differences may have an impact on the better results and higher assessment for learning that were presented in RQ1.

Table 6: Results for process and team variables (RQ2)

RQ2

Scale		N	Mean Rank	Sum of Ranks	Mann-Whitney -U-Test	Sig. (2-tailed)	Effect size r
Student Engagement (ZMS)	Control	65	49,38	3210,00	1065	0,024	0,216
	Treatment	44	63,30	2785,00			
	Total	109					
Team_Atmosphe re (ZMS)	Control	65	46,28	3008,50	863,5	0,000	0,351
	Treatment	45	68,81	3096,50			
	Total	110					
Team_Task Orientation (ZMS)	Control	65	49,62	3225,00	1080	0,020	0,223
	Treatment	45	64,00	2880,00			
	Total	110					
SG_Comprehensi on (ZMS)	Control	65	50,58	3287,50	1142,5	0,050	0,187
	Treatment	45	62,61	2817,50			
	Total	110					
Learnability (MEEGA+)	Control	65	48,99	3184,50	1039,5	0,009	0,248
	Treatment	45	64,90	2920,50			
	Total	110					
Social Interaction (MEEGA+)	Control	65	50,52	3284,00	1139	0,044	0,192
	Treatment	45	62,69	2821,00			
	Total	110					

RQ3: What levels of learning/reflection do we see in the notes on the flips?

In the following we present the results of our qualitative analysis. From gathered answers on the reflection questions we can deduct that students discussed a broad variation of topics. Potential learnings are diverse and evolve during the game play over different game rounds. The teams made clear points on what they thought went well, what they considered not going so well and how they thought they could improve this.

On a regular basis it is visible that goals were achieved and new goals or topics were set every round or previous goals were further specified. These are both indicators for learning, because it is often visible that a preset goal is achieved in the next round or otherwise on the final third flip over. A further specification of a goal can be seen as deepening of insight (Ries, Schaap, van Loon, Kral, & Meijer, 2022). Sometimes it is not visible a goal has been achieved because no references are made toward it, however we do find participants set new goals/topics that could not have been set if they hadn't solved the other issues. The learning aim of the simulation games used in this study is to manage as a team a large variety of indicators via sharing of information, planning and interpretation of results. Therefore, it makes sense the teams continuously add new topics especially when knowing the complexity of the game increases with more information and indicators as the game rounds proceed.

With RQ3 we address the question on what levels the teams reflected during the game. Looking at the four seminars using the sg GM we see a majority of 67% of first order learning (Table 7). This can be logically explained by the goals of the game. Participants should apply their content and procedural knowledge to the game to be successful. Second order learning accounted for 24 percent. This makes sense because regular reflection on the group process can aid quality in decision making and learning from the multiple perspectives in the cooperation. In some instances also third order reflection took place evoked by the reflections on their personal roles in the group and what they needed to contribute to the organization.

Table 7: Levels of learning per simulation game

	TOPSIM GM (8 groups)	Riva STC (2 groups)
type of learning	Percentages (total)	Percentages (total)
first order learning	67% (142)	43% (19)
second order learning	24% (51)	36% (51)
third order learning	9% (18)	21% (18)
Total	100% (211)	100% (44)

For the change management game we see a majority of 43 percent in first order learning (Table 7). This can be explained by the goal of this game, which had less emphasis on applying pre-determined learned norms and rules because the players actually had to find out how to influence the game in such a way they developed an effective approach. Therefore, the process component logically had a more prominent role in this game and accounted for 36% of the cases. While third order learning accounted for another 21%, which makes sense because there was a stronger emphasis on the process and the participants also needed to reflect more on their role in the process and their added value.

A remark has to be made on the sample size which was considerably smaller for STC as opposed to the general management game (eight teams in GM, two teams in STC). Though the sessions were representative to other sessions as confirmed by the experienced facilitators with this game.

What does this mean in relation to the goal of the game? In GM the learning goal for the students was to run a profitable sustainable company. In STC students' learning goal was to move organizational members through the phases of organizational change. The type of learning goal combined with the characteristics of the sg's and the facilitation interventions have implications for the type of learning taking place but how these actually relate and what the exact implications are is unknown. We state this cautious because learning as intended doesn't necessarily deliver learning as happened (Leigh, 2003; Wijse-van Heeswijk et al., in press).

For both games we count a majority of first order statements on the reflection notes (e.g., “investment in new machines” or “not reached our sales target”). It shows us that teams are involved in understanding the content of the simulations and finding good decisions. This is in line with the ideas of the games: they are designed and used to apply knowledge and experience the interdependencies of systems. Seeing also a significant number of second and even third order statements emphasizes that students learning with simulation games are active participants in simulated social organizations. They are not only confronted with content but must make organizational decisions. In the third GM seminar one team reflected on the exchange of information within the team. This second order statement can be seen as an example for organizational decisions teams had to make. It can be assumed that decisions on different levels (first, second, third order) interact with each other. The organization of the exchange of information within a team may have great impact on facts that are available for decision making. Having not enough facts (first order) for decision making could lead to the idea to change the organization within the team (second order).

RQ4: What learning process do we see in the teams from round to round?

The reflection-notes on the flips document the group discussion in a chronological manner. So, the notes provide insight into a process of discussion and learning the treatment groups went through. As an example, two out of ten group reflections are elaborated in detail here. For these two examples all data can be found in the appendix (more data are available with the researchers). Next to the examples a summary containing data from all groups is given.

Example 1: GM1 Team A (see appendix; at the end of [this](#) manuscript):

For this group most of the topics occur only once (24 times, scenario 1). It remains unclear why the groups lost track of the topic. Possibly students learned immediately how to deal with the topic or found the topic irrelevant in later rounds. For scenario 1 it is unclear. We count three issues that were discussed as problematic in reflection one and two (“Calculation of cost of goods sold for copy budget” (1) “Staff utilization too high” (2), “Sales volume estimated too high” (3)) but were discussed as “went well” in reflection 3, which is labeled as scenario 3. In

reflection one and two the topic of appropriate staff utilization was discussed as problematic (scenario 4). Only one topic was discussed as satisfying in the beginning (“good share price”) but as problematic in later rounds, which therefore is labeled as scenario 5.

Example 2: STC Team blue (see appendix; at the end of [this](#) manuscript)

For this team we see seven examples of scenario 1. Scenarios 2 to 4 are found once each and no example of scenario 5 is seen. After round one the team stated that `a plan` was needed. Indeed, it is essential for the simulation to bring 42 action cards into a structured and continuing order. After identifying their lack of a plan, the team obviously found a functional pattern (a plan) because it was mentioned in all later reflections as positive. After round one the team identified a major problem and was able to work on it in the next rounds.

These two examples lead to an inductive hypothesis explaining why we see a better game performance for treatment groups and higher values for self-reported learning: The structured reflection assignment helps students to identify and formulate relevant problems to work on. Formulating issues may reduce complexity and help teams to focus on relevant problems to work on. This hypothesis is confirmed by summarizing the qualitative data. As already seen in the described examples the first scenario is found most for all ten groups (Table 5). Due to the open reflection questions, it is logical to see topics that occur only once or are only relevant for one round. Learners proceed in learning and things that are learned (or do not matter anymore) can disappear from the reflection list. In 20 instances we see teams identifying an issue as problematic but finding ways to deal with it in a positive way in later rounds (scenario three). Identifying an issue as positive in more than one round (scenario two) is found 11 times. Altogether 15 times we see teams naming topics as problematic in more than one reflection round, indicating that teams were challenged by the topic and did not find satisfying solutions for them yet (scenario four). In only three cases we see issues `falling` from positive in earlier rounds to problematic in later rounds.

Looking only at the four scenarios (two to five) that afford a topic to be mentioned at least in two reflections we can state that scenario three appears most frequent (Table ten following). It confirms the inductive thesis that the consecutive reflection

assignment helps groups to identify relevant topics and to work on them successfully.

Table 10: Summary Learning process

Scenario	Count	Total	
Topic occurs only once	Scenario 1	132	132
positive -> positive	Scenario 2	11	31
negative -> positive	Scenario 3	20	
negative -> negative	Scenario 4	12	15
positive -> negative	Scenario 5	3	

Conclusion and Discussion

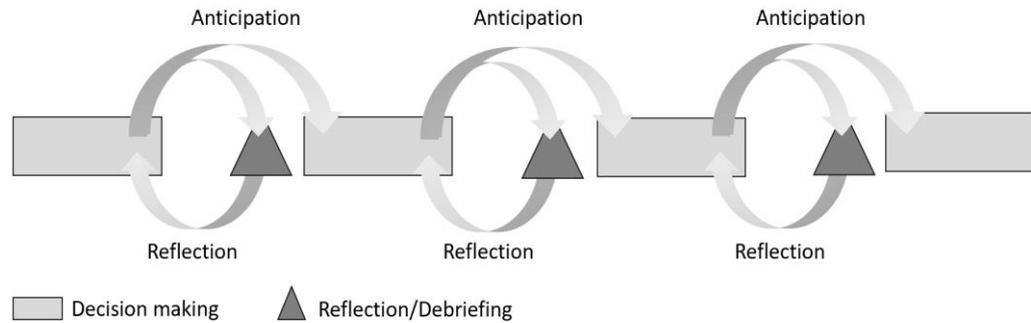
With a quasi-experimental research design the role of pre structured consecutive reflection questions for learning in sg was researched. In five sg-based seminars treatment groups participated in three reflection interventions (start, middle, end) with three pre structured open questions while non-treatment groups did not. It was found that treatment groups performed better in the games and assessed their self-reported learning higher when asked for concrete learning goals (RQ1). Differences between treatment and non-treatment students are also seen in the assessment of the gameplay process (RQ2). Effect sizes for sg-results and learning (RQ1) and engagement, teamwork and sg-comprehension (RQ2) were small and medium. With regard to the scope of the intervention this is still a remarkable result. In block seminars lasting two to three days three reflection sessions of fifteen minutes each made a difference. Treatment students see themselves to be more engaged in the game and have higher values for teamwork (ZMS inventory) and social interaction (MEEGA+). Also, the learnability (MEEGA+) and the comprehension (ZMS inventory) of the game was rated higher by treatment students, indicating that they had a better understanding of the sgs. Analyzing the reflection notes of treatment groups with qualitative methods we mainly find examples for first order learning and a tendency of more second and

third order learning in the change management simulation. It can be assumed that the simulations address different levels of learning (RQ3). Looking at the reflection notes in a chronological order we find reflection-groups to identify relevant problems and to manage them successfully (RQ4).

These results provide a congruent picture. The reflection intervention required group discussion and the results emphasize that treatment students rated their group experience and social interaction higher than non-treatment students. Treatment students rated themselves to have a slightly better understanding of the simulation games which is in line with better self-reported learning and even better success values by the end of the simulation games. Structured group discussion with open reflection questions (What went well?, What went not so well? and What do we need to improve?) may have led to a better understanding of the relevant problems for the simulation which leads to better results. In our qualitative analysis we see further indications explaining why treatment teams have better results and might have profited more from the seminars. The repeated reflection assignment helped teams to identify relevant problems and in many cases the teams were able to work on them properly (RQ4). For GM we reported roundly results (for STC roundly results are not available) and it is obvious that only in the last rounds treatment and non-treatment teams differentiated. This matches with our qualitative results. The reflection assignment serves as a circular learning mechanism. It takes repetition identifying issues or goals, working on them, reflecting them again and then to restart the circle. This is in line and confirms the theory on experiential learning. Experience and reflection complement each other in a circular relationship (Hilzensauer, 2008; Kolb & Kolb, 2009). Interrupting the gameplay with reflection assignments gives students the chance to look back to decision making and to identify relevant issues as shown in graphic 2. By this process treatment teams gained focus in a complex and confusing environment. The relevant issues found are not only reflected but also anticipated for the next round of decision making. This leads to clearer ideas for action in the next rounds and provides players with agency. We think it is mainly this permanent circle of reflection and anticipation that helped treatment teams to understand the games better and therefore to achieve better results.

The overall conclusion of this study is adding structured reflection questions for (in-between-) debriefings during sg sessions adds value to learning.

Graphic 2: Process of decision making and reflection



Limits and further recommendations

Even though the results are congruent and plausible they should be confirmed with further research and more participants/groups. Students participating in this study are all business students from one University in Stuttgart, Germany. To confirm and generalize these results more research is needed applying different contexts, cultures and professions.

We were able to use at least two different simulation games in this study. Further research should evaluate if the presented results are confirmed when applying the reflection assignment to other simulation games.

The discussion notes that were used for qualitative analysis represent only the results of the discussion and give only a narrow insight into the discussion. To understand this even better a deeper insight is needed and further qualitative data of group discussions and gameplay (voice, video) could be collected and analyzed. Especially the reflection notes of one group were almost the same from round to round and researchers were not sure how seriously the group took the discussion.

References

- Alklind Taylor, A.-S. (2014). Facilitation matters: A framework for instructor-led serious gaming. Skövde: University of Skövde.
- Argyris, C. (1977). Double Loop Learning in Organizations. Retrieved from <https://hbsp.harvard.edu/product/77502-PDF-ENG>

- Argyris, C., & Schon, D. A. (1974). *Theory In Practice: Increasing Professional Effectiveness*. San Fransisco: Jossey-Bass.
- Baker, A. C., Jensen, P. J., & Kolb, D. A. (1997). In Conversation: Transforming Experience into Learning. *Simulation & Gaming*, 28(1), 6–12.
<https://doi.org/10.1177/1046878197281002>
- Bartunek, J. M. (2014). From the Editors: Honoring the Legacy of Chris Argyris by Devoting Attention to How Managers Learn. *Academy of Management Learning & Education*, 13(1), 1–4. <https://doi.org/10.5465/amle.2013.0369>
- Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a Taxonomy Linking Game Attributes to Learning. *Simulation & Gaming*, 43(6), 729–760. <https://doi.org/10.1177/1046878112439444>
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-12770-0>
- Brazhkin, V., & Zimmerman, H. (2019). Students' Perceptions of Learning in an Online Multiround Business Simulation Game: What Can We Learn from Them? *Decision Sciences Journal of Innovative Education*, 17(4), 363–386.
- Cloutier, C., & Langley, A. (2020). What Makes a Process Theoretical Contribution? *Organization Theory*, 1(1). <https://doi.org/10.1177/2631787720902473>
- Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44(04), 823–830. <https://doi.org/10.1017/S1049096511001429>
- Crossan, M. (2003). Chris Argyris and Donald Schön's Organizational Learning : There is no silver bullet. *Academy of Management Perspectives*, 17(2), 38–39. <https://doi.org/10.5465/ame.2003.10025187>
- Deen, M. (2015). *G.A.M.E. Games Autonomy Motivation & Education*. Eindhoven University of Technology, Eindhoven.
- Duke, R. D. (2014). *Gaming: The future's language*. Bielefeld: W. Bertelsmann Verlag.
- Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statisitk und Forschungsmethoden*. Basel: Belz Verlag.
- Friedman, V. J., Torbert, W. R., Nielsen, R., Silverman, H., & Bradbury, H. (2014). Chris Argyris (1923–2013): In memoriam. *Action Research*, 12(1), 110–113. <https://doi.org/10.1177/1476750313519542>
- Geithner, S., & Menzel, D. (2016). Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. *Simulation & Gaming* 47. (2), 228–256. Retrieved from <http://nbn-resolving.org/urn:nbn:de:bsz:14-gucosa2-353942>
- Geurts, J. L., Duke, R. D., & Vermeulen, P. A. (2007). Policy Gaming for Strategy and Change. *Long Range Planning*, 40(6), 535–558.
<https://doi.org/10.1016/j.lrp.2007.07.004>
- Greenwood, J. (1988). The role of reflecion in single and double loop learning. *Journal of Advanced Nursing*. (27), 1048–1053.
- Hilzensauer, W. (2008). Theoretische Zugänge und Methoden zur Reflexion des Lernens. Ein Diskussionsbeitrag. *Bildungsforschung*, 5(2). Retrieved from <http://www.bildungsforschung.org/Archiv/2008-02/lernvermoegen/>

- Hühn, C., & Rausch, A. (2022). Collaboration and emotions during simulation-based learning in general management courses. *Studies in Educational Evaluation*, 73, 101130. <https://doi.org/10.1016/j.stueduc.2022.101130>
- Husebø, S. E., Dieckmann, P., Rystedt, H., Søreide, E., & Friberg, F. (2013). The relationship between facilitators' questions and the level of reflection in postsimulation debriefing. *Simulation in Healthcare : Journal of the Society for Simulation in Healthcare*, 8(3), 135–142. <https://doi.org/10.1097/SIH.0b013e31827cbb5c>
- Klabbers, J. H. G. (2009). *The magic circle: Principles of gaming & simulation* (3rd and rev. ed.). Rotterdam: Sense Publishers.
- Kolb, A. Y., & Kolb, D. A. (2009). The Learning Way. *Simulation & Gaming*, 40(3), 297–327. <https://doi.org/10.1177/1046878108325713>
- Langley, A. (1999). Strategies for Theorizing from Process Data. *Academy of Management Review*, 24(4), 691. <https://doi.org/10.2307/259349>
- Langley, A., & Tsoukas, H. (2016). *The SAGE Handbook of Process Organization Studies*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd. <https://doi.org/10.4135/9781473957954>
- Lee, J. Y., Donkers, J., Jarodzka, H., Sellenraad, G., & van Merriënboer, J. J. (2020). Different effects of pausing on cognitive load in a medical simulation game. *Computers in Human Behavior*, 110, 106385. <https://doi.org/10.1016/j.chb.2020.106385>
- Leigh, E. E. (2003). *A practitioner researcher perspective on facilitating an open, infinite, chaotic simulation : learning to engage in theory while putting myself into practice* (Thesis). University of Technology, Sydney. Retrieved from <http://hdl.handle.net/10453/20068>
- Muno, W. (2016). Fallstudien und Process Tracing in der Vergleichenden Politikwissenschaft. In H.-J. Lauth, M. Kneuer, & G. Pickel (Eds.), *Handbuch Vergleichende Politikwissenschaft* (pp. 79–90). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-02338-6_6
- Nakamura, M. (2022). The Effects of Structured Instruction on Team Performance. In U. Dhar, J. Dubey, V. Dumblekar, S. Meijer, & H. Lukosch (Eds.), *Gaming, Simulation and Innovations: Challenges and Opportunities* (pp. 15–27). Cham: Springer International Publishing.
- Pasin, F., & Giroux, H. (2011). The Impact of a Simulation Game on Operations Management Education. *Computers & Education*, 57(1), 1240–1254.
- Petri, G., Gresse von Wangenheim, C., & Borgatto, A. F. (2017). MEEGA+, Systematic Model to Evaluate Educational Games. In N. Lee (Ed.), *Encyclopedia of Computer Graphics and Games* (pp. 1–7). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-08234-9_214-1
- Petri, G., Gresse von Wangenheim, C., & Borgatto, A. F. (2018). MEEGA+: A Method for the Evaluation of Educational Games for Computing Education. Universidade Federal De Santa Catarina, Brazil.
- Pospeschill, M. (2022). *Testtheorie, Testkonstruktion, Testevaluation* (2nd ed.). München: Ernst Reinhardt Verlag.
- Reynolds, M. (2014). Triple-loop learning and conversing with reality. *Kybernetes*, 43(9/10), 1381–1391. <https://doi.org/10.1108/K-07-2014-0158>

- Ries, K. E. de, Schaap, H., van Loon, A.-M. M. J. A. P., Kral, M. M. H., & Meijer, P. C. (2022). A literature review of open-ended concept maps as a research instrument to study knowledge and learning. *Quality & Quantity*, 56(1), 73–107. <https://doi.org/10.1007/s11135-021-01113-x>
- Riva solutions GmbH. SysTeamsChange®: Das führende Planspiel zum Veränderungsmanagement. Retrieved from <https://www.systems-change.com/>
- Schwägele, S., Zürn, B., Lukosch, H. K., & Freese, M. (2021). Design of an Impulse-Debriefing-Spiral for Simulation Game Facilitation. *Simulation & Gaming*, 52(3), 364–365. <https://doi.org/10.1177/10468781211006752>
- Stoppelenburg, A., Caluwé, L. de, & Geurts, J. (2012). *Gaming: Organisatieverandering met spelsimulaties (geh. herz. editie)*. Deventer: Kluwer.
- Taylor, A.-S. A., Backlund, P., & Niklasson, L. (2012). The Coaching Cycle. *Simulation & Gaming*, 43(5), 648–672. <https://doi.org/10.1177/1046878112439442>
- TOPSIM GmbH. TOPSIM - General Management. Retrieved from https://www.topsim.com/wp-content/uploads/2021/03/SD_TOPSIM_General-Management.pdf
- Tosey, P., Visser, M., & Saunders, M. N. K. (2012). The origins and conceptualizations of 'triple-loop' learning: A critical review. *Management Learning*, 43(3), 291–307. <https://doi.org/10.1177/1350507611426239>
- Trautwein, F., & Alf, T. (2022). Theoriebasierte Entwicklung eines Inventars zur Evaluation von Planspielveranstaltungen. In T. Alf, S. Hahn, B. Zürn, & F. Trautwein (Eds.), *Planspiele - Erkenntnisse aus Praxis und Forschung: Rückblick auf den Deutschen Planspielpreis und das Europäische Planspielforum 2021* (pp. 63–87). [S.I.]: Books on Demand.
- Visser, M. (2007). Deutero-Learning in Organizations: A Review and a Reformation. *Academy of Management Review*, 32(2), 659–667. <https://doi.org/10.5465/AMR.2007.24351883>
- Visser, M., Chiva, R., & Tosey, P. (2018). Guest editorial. *The Learning Organization*, 25(4), 218–223. <https://doi.org/10.1108/tlo-02-2018-0021>
- Watt, K., & Smith, T. (2021). Research-Based Game Design for Serious Games. *Simulation & Gaming*, 52(5), 601–613. <https://doi.org/10.1177/10468781211006758>
- de Wijse-van Heeswijk, M., Rouwette, E., van Laere, J. (2023). Case Study Report on Facilitation Interventions to Increase Learning Effectiveness in Game Simulations. In: Angelini, M.L., Muñoz, R. (eds) *Simulation for Participatory Education*. Springer Texts in Education. Springer, Cham. https://doi.org/10.1007/978-3-031-21011-2_5

Appendix

Coding Scheme: Levels of Learning for RQ3

Developed by Marieke de Wijse-Van Heeswijk MSc, Prof. Etienne Rouwette and Prof. Sander Meijerink in June 2021

<p>Code 1 participant response written on flap</p>	<p>1. PARTICIPANT FIRST ORDER (CONTENT/PROCEDURAL) REFLECTION OR RESPONSE</p> <p><i>This type of reflection or response concerns first order behaviour, such as dealing with norms (but nor questioning or changing them!), how or when someone applies a norm, or is about applying procedures based on norms are expressed, how or when someone applies a procedure. (Argyris 1983, Argyris 2002, Gugerell and Zuidema 2017, Fazey, Schöpke et al. 2018).</i></p>	<p>Examples</p>
	<p>This code captures either a content and/or procedural response or learning. This can be an application of an indicator or procedure in the game. The relevance of the indicator is recognized by the participants when it is put on any of the questions on the flaps.</p>	<p>Examples of the content response: “The participant says: “Yes. I have the incidents here, I still have to prioritize them, but I am not sure how many incidents are still in the system outside of my view”.</p> <p>The participant notes on a flap Absatzplanung, or Forschung erhöhen.</p> <p>”</p>

		<p>Examples of the procedural response:</p> <p>Participant writes down: follow application procedure</p> <p>The participant says: "you have to pass the form to me according to the procedure"</p> <p>The participant says: "Now I know how I can respond to these incidents I have to consult the coordination department."</p>
<p>Code 2 participant response written on flap</p>	<p>2. PARTICIPANT (PROCESS) SECOND ORDER REFLECTION OR RESPONSE</p> <p><i>Participants that discuss, reflect and/or change their assumptions and norms on processes and are engaged in (parts of) double loop learning processes.</i></p>	<p>Examples</p>
	<p>Process reflections entail investigating the basic values and emotions of the people involved. It concerns the process of participants of making sense and reflecting on their functioning, the possible mismatch between actions and norms, and the adaptation to a more functional or contributing process. Double loop learning is needed when the original norms and assumptions behind the norm require change (Argyris 1976, Reynolds 2014).</p> <p>Second order learning description: Participants are explicit in changing norms or discussing norms and making adaptations to their previous behavior in dealing with</p>	<p>The participant says: "Would it not be better to coordinate actions instead of just being in the control modus signalling mistakes?"</p> <p>The participant says: " In the coordinative role I can contribute to the processes preceding the occurrence of errors."</p> <p>The participant writes down:</p> <p>"Kommunikation war gut"</p> <p>"Gute teamarbeit"</p> <p>"Effiziente entscheidungen"</p>

	<p>disturbances of any kind. These disturbances can be characterized as challenges to their norms and underlying values. Targets sometimes are reframed and adapted to the actual conditions together with the changes in processes and internal infrastructure (Argyris 1983, Argyris 2002, Gugerell and Zuidema 2017, Fazey, Schöpke et al. 2018). The participants can for example decide to redesign their processes and try out new ways of working. Experimentation on a process level belongs to this category.</p> <p>Participants can also make a remark on how they have experienced certain factors in the game such as experienced time in the game, or group cooperation or a remark on the decision making process.</p>	
<p>Code 3 Participant response written on flap</p>	<p>3. PARTICIPANT THIRD ORDER REFLECTION/RESPONSE OR LEARNING</p> <p><i>Participants show a meta reflection or response to challenges from their role and/or personal perspective on how they perceive their role and its added value or missing added value. For example also what they perceive as needs they require from their roles.</i></p>	<p>Examples</p>
	<p><i>Participants respond with a perspective on their role in the game or outside the game, how do they add value for themselves and for others from their role? (it can also be a general</i></p>	<p>The participant says: “ “I am quitting this game I am not putting up with this mess as director any longer, I have had it. I cannot add value from my role because I am</p>

	<p><i>statement on roles and role taking in the simulation game)</i></p> <p>Participant(s) respond with a reflection and or learning statement on their personal role and or roles of others on how they deal or want to deal with the role. Sometimes they do not make explicit in their response what they have learned although it is likely third order learning is present.</p> <p>Participants reflect on how they see themselves in their roles and how they perceive the added value for others in their role. They investigate the assumptions and hidden rules behind behavior in the light of the functioning of the organization and its viability (Visser 2007). This way of reflection leads to learning to learn, to look at functioning from different levels and perspectives and understanding the implications (Reynolds 2014). This type of learning can lead to transformational learning (Kuntz and Gomes 2012) and can lead to fundamental change in roles and behaviors within and outside social systems (Tosey, Visser et al. 2012).</p> <p>Participants use metaphors to relate to their situation in the gameplay and their reality. They draw relations to how processes in the gameplay relate to their reality and how they perceive this as either functional, neutral or dysfunctional.</p>	<p>restricted by the game rules and I feel I need more autonomy also in my real professional life.”</p> <p>The participant writes wir sind nicht genug positive gewesen, solten denken in moglichkeiten und wachsung</p> <p>The participant says: “I feel I am overloaded with issues I have no authority of just like in practice. I will go to the manager and propose I am informed in another way and I get more autonomy to deal with issues I can fix myself”.</p> <p>The participant says: “in future I will view my role as a colleague differently. I will not be just responsive and pass-through things I do not understand, but I will take up an active and learning role investigating matter I do not directly understand or comprehend and share my perspectives with my coworkers in order to learn and attain a more proactive role”.</p> <p>The participant says: “as a director it is challenging to stay in contact with the process”</p> <p>The participant says: “I think we play the idealistic world here, this is not how It goes in our reality”</p>
--	--	--

	If the remark about role is related to a prescribed task then it is not a role reflection because it is not a possibility to adapt norms of take on a specific role, it is about a procedure.	
Code tangent 't' For not codable units	<i>Tangent for text that cannot be coded in one of the categories, for instance for coders who were not present at the session or for instance for bits of text that do not fall in the category of the codes</i> <i>Unless the answer is directly related to a (dis)confirmation to a remark or question right before the code</i>	The facilitator nods The facilitator says 'yes' The facilitator says 'no' The facilitator says 'ok' The facilitator says 'good' <i>"the facilitator pauses"</i> <i>"the participant says I think... (and does not further explain)"</i>

Examples of Flip-Reflection-Notes (RQ3)

The following table shows with what colors the scenarios are marked. If a scenario occurs more than once, it is also marked with a number to identify the issue in a later round.

Topic occurs only once	Scenario 1
positive -> positive	Scenario 2
negative -> positive	Scenario 3
negative -> negative	Scenario 4
positive -> negative	Scenario 5

Reflection Notes GM1 Team A (own translation from German to English)

reflection flap 1	Scenario	reflection flap 2	Scenario	reflection flap 3	Scenario
What went well?		What went well?		What went well?	
Communication		Investors Presentation		Estimating Staff Utilization	2
Consultation before decision		Continued positive EBIT		Estimation of sales & credit	3
Courage for decision making		Planning well thought out		low cost of goods sold	1
Branding		sound investments		empty stock (budget)	
Good results high revenue				Payout of dividends (?)	
good share price					
low debt					
What went not so well?		What went not so well?		What went not so well?	
Missverständnisse		Staff utilization too high	2		
übersehene Aspekte (Fluktuation)		Sales volume estimated too high	3		
What do we need from ourselves and others to reach our goals?		What do we need from ourselves and others to reach our goals?		What do we need from ourselves and others to reach our goals?	
Calculation of cost of goods sold for copy budget	1	Gen 2 - Market Dominance		Higher staff utilization	
Cashflow betrachten		make market 2 more profitable		Empty warehouse	
		new machines		Increase of the share price	
		more FK 70/30			
		Leasing			

Reflection Notes STC1 Team blue (own translation from German to English)

reflection flap 1	Scenario	reflection flap 2	Scenario	reflection flap 3	Scenario
What went well? Team, Communication Decisions ->Information slows down decisions	1	What went well? Planning Communication Teamwork	1 1	What went well? Plan Communication People in integration Industrial espionage	1
What went not so well? Not thought through to the end card No. 18, No. 8		What went not so well? no 4th ring Waiting time is too long, costs Euros		What went not so well? Cards partly unclear Credit needed	
What do we need from ourselves and others to reach our goals? a plan Having an overview of the prerequisite of cards		What do we need from ourselves and others to reach our goals? knowing which cards give the rings		What do we need from ourselves and others to reach our goals? better understanding of the cards	

5.4 Impact Factors on Learning and Satisfaction in Simulation Game Based Courses in Higher Education

Abstract

Simulation games (SG) are complex interactive learning environments where multiple aspects interact with each other and contribute to the learning process. The aim of this research is to evaluate which factors influence learning and satisfaction in courses that are taught with SGs. Using Partial Least Square-Structural Equation Modelling this study analyzes several dimensions of teaching and learning with SGs based on a theoretical model. The study uses evaluation data from 192 SG-based courses (n=3067) using 32 different SGs. The research reveals the highest direct impact on learning and satisfaction from the simulation game, followed by facilitation. Putting direct and indirect effects together facilitation has the strongest effect on learning and satisfaction in SG-based courses. This means, learning in the context of simulation games is foremost influenced by facilitators, but mediated by the simulation game. In addition, but to a lesser extent, student engagement influences learning and satisfaction as well as the understanding of simulation games. Contrary to expectations, teamwork hardly influences learning and satisfaction.

Autoren: Tobias Alf, Friedrich Trautwein

Introduction

During a conference with many experts on the spot Hofstede et al. (2010) raised the question why games work for learning. Interviewing about 20 experts in the field during the conference they tried to find the “active substance” that is at work when using simulation games (SG). It is not surprising that they concluded “that an enormous variety in aspects is stressed by the 20-odd practitioners who participated. Gaming simulation is multifaceted. No single schema can do justice to the range of aims that can be achieved by simulation gaming.” (p. 839) The aim of this article is similar. Using empirical student evaluation data, we try to find the “active substance”. Hofstede et al. followed a qualitative approach which broad a very wide field of “active substances” to the surface that they clustered in knowledge, social aspects, emotions, and practice on the one hand and in design, play and debriefing on the other hand (p. 840). Since we are using a quantitative approach, we have the disadvantage that we cannot consider such a rich field of “active substances” that might be at work and cannot inductively create new variables/categories. But using quantitative data we have the advantage that we can model direct and indirect effects in a Structural Equation Model. We therefore can explore how the variables interact and possibly moderate each other. Probably we also will not find *the* active substance, but how various “substances” play together.

Data for this research was gathered using a questionnaire that was developed to evaluate the use of SGs in higher education courses (F. Trautwein & Alf, 2023). It is inspired by the evaluation of teaching model according to Rindermann (2009). In this study we therefore take three groups of independent variables into account: 1. facilitation, 2. students and 3. simulation games. All three directly impact learning and satisfaction but also interact with each other. Therefore, we raise four research questions:

- RQ1: How do students affect learning and satisfaction?
- RQ2: How do facilitators affect learning and satisfaction?
- RQ3: How do simulation games affect learning and satisfaction?
- RQ4: How do students, facilitators and simulation games interact with each other and affect learning and satisfaction?

In the context of this study, we understand SGs as educational environments in the spectrum of roleplay, game and simulation without being fully attributable to

one of these dimensions (Kriz & Manahl, 2016). Students take over roles and act/play/decide according to rules in simulated environments using resources (Klabbers, 2018). The use of SGs in educational contexts is often related to problem-based learning (Mandl, 2004), experiential learning (Geithner & Menzel, 2016) and constructivist learning theory (Kriz, 2010) in general. One great benefit of the use of SGs is the integration of cognition, emotion and action (Hofstede et al., 2010). They are multidimensional learning environments.

In the following Literature Review we describe all variables taken into consideration for this research and describe theory-based how they might affect each other. The paragraph Data and Methods explains how the data was gathered and how the data was analyzed to answer the research questions raised. The paragraphs Results and Discussion are structured according to the research questions and describe and discuss the results found. In a Conclusion we sum this research up and outline practical and theoretical implications.

Literature Review

In this research we explore the impact of students, teachers/facilitators and SGs on students' learning and satisfaction. First, we introduce the dependent variable learning and satisfaction, followed by the three impact factors and how they might interact with each other.

The dependent variable: Learning-and-Satisfaction

The main goal of educational activities is learning. But learning can mean a very broad field of things: The acquisition of conceptual knowledge, process knowledge, tactical knowledge, social, personal, or vocational competencies to name only some. SGs are indeed used for a very broad field of learning goals and a whole variety of subjects (Alf, 2022). Quite often, they are designed to teach dealing with complexity and decision making (Pasin & Giroux, 2011) or gain systems insight (van Blisen et al., 2010). Many studies and literature reviews have shown that SGs can be powerful and holistic educational methods (Zeiner-Fink, 2023). Next to the focus on certain learning goals it should be considered that education is a long process of sequential experiences. Following John Dewey's foundations of experiential education, the concrete experience of an educational activity has a

huge impact not only for the certain activity but also for those to come. The experience in one activity potentially affects the attitude and experience of further educational activities and so the whole educational attitude of a person (“experiential continuum”) (1938, p. 33). Therefore, we understand learning and satisfaction in a broader way than only focused on specific learning goals. Following the idea of the experiential continuum we also emphasize that the quality of the experience (satisfaction) is very relevant.

Therefore, the aim of this research is to evaluate impact factors based on a wide variety of SGs and not to focus on certain learning goals. Using the questionnaire from F. Trautwein and Alf (2023), we therefore measure learning by self-reports. It contains three items that focus more on learning and three that relate more on satisfaction. All items used for this study are described in the paragraph Data and Methods (Table 2).

Students and their impact on learning and satisfaction (RQ 1)

In constructivist learning theories, learners are described as active and responsible for the learning process. Especially in research on learning with SGs it is emphasized that students affect their own learning success by active participation. Mayer et al. (2013) found effects from motivation, attitude towards game based learning and engagement in the game on learning-satisfaction. But they emphasize that it is especially the effort/engagement students put in the game that determines learning-satisfaction. Corbeil and Laveault (2011) investigate effects of a SG on history teaching and also found significant correlations between students’ involvement and learning (measured with test results) while other factors such as motivation or interest on the topic were of minor relevance. Tao et al. (2012, p. 1357) found corresponding results: Students that engaged more in group discussion during a SG had a better game performance. Alas et al. (2018) researched factors that influence success in business games. They found significant positive correlations between student’s engagement and success in the game. Students that prepared homework on the simulation were more successful and students who dealt more with information provided by the game were also more successful. On the other hand, Alas et al. found significant negative correlations between absence during the SG and success. Overall, the referred studies point into the same direction: With their own activity and engagement

during the SG students have an impact on their own learning and/or success in the game.

Typically, students participate in SGs not individually but in teams. Using SGs in education students do not only work on the content of the game, but also deal with the challenges of teamwork (finding roles and responsibilities, communication) (Brazhkin & Zimmerman, 2019, p. 369). Several studies show significant correlations between successful teamwork and learning respectively success in SGs. Van den Bossche et al. (2011) for example researched to what extent shared mental models within teams support the teams' performance. They find positive correlations between shared mental concepts and different success indicators in the SG. This is also supported by the research of Hwang (2018) who also explores relationships between teamwork and performance in SGs. We can conclude that there is a correlation between the way students work together in teams and their performance in the game. It should be noted that all studies (known to the authors) in this context have explored a relationship between teamwork and game performance. Sometimes performance is used as an indicator for learning, but it is not necessarily learning. In the context of this study, we do not use game performance measurement and therefore evaluate the relationship between teamwork and self-assessed learning and satisfaction.

Based on the questionnaire of F. Trautwein and Alf (2023) student engagement is measured with two items that are combined to one scale. One item measures activity during decision/working phases in teams and another item scores activity in general. A third item on student engagement, as suggested by Trautwein and Alf, is not included in this research. This is due to insufficient reliability. Teamwork is a very complex construct that can be differentiated in various dimensions (Hwang, 2018). In the context of RQ1 we focus on the impact students have on learning and satisfaction. In the empirical analysis we therefore focus on items that evaluate the function and tasks students had within the team (F. Trautwein & Alf 2023).

Based on the presented theory we evaluate RQ 1 (How do students affect learning and satisfaction?) by implementing the following hypotheses into the model:

- Student engagement has a positive effect on learning and satisfaction.
- Students function and task within the team has a positive effect on learning and satisfaction.

Facilitators and their impact on learning and satisfaction (RQ2)

Learning games as described by Plass et al. (2015) are designed to individually play a game and thereby learn. This is not exactly the case for SGs. Of course, students can learn in SGs during their playing activities, but SGs are introduced and guided by teachers. Therefore, their impact needs to be considered when evaluating impact factors on learning and satisfaction in SG-based teaching and learning. But the role of teachers when using SGs is different from the role of teachers in traditional teaching. When using SGs in educational contexts lecturers become facilitators. They don't teach but guide and support students on their own learning path. Facilitation can be "conceptualized as activities a facilitator can predesign or facilitation during gameplay and reflection aimed at enhancing learning from the simulation game" (Wijse-van Heeswijk, 2023, p. 54). Facilitation means a broad bundle of attitudes, knowledge and skills lecturers should have and apply when guiding groups in their educational activities (Kortmann & Peters, 2021, p. 261). Kortmann and Peters (2021) describe facilitators as "unseen helmsman", meaning that facilitators guide the activity in a way that students can make their own self-directed learning experience. Facilitators remain hidden ("unseen") but still guiding the activity. Closely related to the idea of facilitation is the concept of debriefing. The experiences and knowledge gathered during the game is made "explicit and shared in the group of players" (p. 257). Not only conceptual papers stress the importance of facilitation. Also, qualitative and quantitative research on the role of facilitators in education with SGs reveals the impact of facilitation on learning. Dube (2015, p. 37) found in her study that the management game she explored was not self-explanatory and guided debriefing-discussions were necessary for students to learn from the game. In accordance to that Richards and Camuso (2015) also stress the importance of individual and collective guided debriefing for learning with SG. Also in quantitative studies teacher related variables are found to be significant predictors for learning. Hense et al. (2009, p. 127) found that teaching experience and motivation are significant predictors for learning with SG. C. Trautwein (2011, p. 220) also finds that teacher related variables are significant predictors. But compared to the importance the topic facilitation has in qualitative research and in conceptual papers the statistical findings are surprisingly low. In none of the models it was the most important predictor.

In the study presented in this paper, facilitation is measured by five items that cover how well the facilitator knows the game, structures the course and how communication with the facilitator was perceived by students.

Based on the described theory to evaluate RQ2 (How do Facilitators affect learning and satisfaction?) we implement the following hypothesis into the model:

- Facilitation has a positive effect on learning and satisfaction.

Simulation Games and their impact on satisfaction and learning (RQ3)

The whole idea about the educational use of SGs is that there is a playful simulated environment in which students gather experiences that are reflected afterwards. So, we can assume that the game, its quality and mechanisms, have a huge impact on satisfaction and learning.

It is a major limitation of almost all empirical studies on SGs that their data is based on only one specific SG. The extensive literature review of Zeiner-Fink (2022) reveals that only five out of 133 studies included in her review look at more than one SG. Since it is our aim to research effects that are valid for educative activities with SGs in general we include various SGs in the study (see Data and Methods).

Referring to Klabbers (2018) SGs can be described as temporal social organizations that consist of three elemental blocks: Actors/Roles, Rules and Resources. Students take over roles and act within the game according to certain rules. According to their roles and the rules of the SG they can use simulated limited resources to reach certain goals within the game. If all players act accordingly, they establish a temporary social system with its own logic (players for example are convinced that a certain card has more power than another or that green plastic chips are more valuable than blue ones). But which attributes should SGs have to be useful mediums for learning? In this context the fidelity or realism of a SG is often found relevant for learning (C. Trautwein, 2011, p. 219; Zeiner-Fink et al., 2023, p. 44). Fidelity means the degree to which a SG represents the reality it is supposed to simulate. But learning with SGs it is not only about the accurate representation of reality. Game designers distinguish between rather abstract models and models that are close to the complexity of reality (cone of abstraction) (Duke & Geurts, 2004, p. 274; Kriz & Manahl, 2016, p. 383). While abstract modelled games can be useful for novices and a broader field of

participants very realistic modelled games tend to be designed for experts in a certain field. So, the very high degree of fidelity in games that are close to reality could also be overcomplex for students and hinder learning. Therefore, SGs should not necessarily have the maximum level of fidelity but be appropriate for the audience and the educational context. Hense et al. (2009, p. 127) found in a regression model that the game-appropriateness was the most important predictor next to 15 other predictors for learning.

The questionnaire of F. Trautwein and Alf (2023) takes two aspects of SGs into account. First the practical relevance of the game. With two items the questionnaire evaluates how well a SG represents reality respectively the practical matter it is designed for. This represents basically what is called fidelity/reality in other studies but takes also in account that the fidelity needs to be on a level that is relevant for learners (SG practical relevance). Second the questionnaire evaluates with three items how well the SG can be learnt and understood by students using it (SG Comprehension).

Based on the presented theory to evaluate RQ3 (How do simulation games affect learning and satisfaction?) we implement the following hypotheses into the model:

- The comprehension of SGs has a positive effect on learning and satisfaction.
- The practical relevance of SGs has a positive effect on learning and satisfaction.

The interplay of Students, Facilitators and Simulation Games (RQ4)

So far students, facilitation and SGs were only described as impact factors (independent variables) that affect learning and satisfaction. It can be assumed that reality is more complex and the so far separated impact factors might also interact with each other. In the development of hypotheses describing relationships between the (actually) independent variables we can draw on previous studies only to a limited extend. The reason is that educational studies usually evaluate learning as their dependent variable and do usually not analyze if teamwork for example affects fidelity. To develop useful hypothesis, we regard to a distinction that was introduced by Corbeil and Laveault (2011, p. 466): Learning about the game and learning from the game. The main idea is that students first must learn about the game. They must understand the scenarios, rules, roles etc. so that they are able

to play the game. But it is not the educational goal to make students play. Enabling students playing a game is a mediator for the actual goal to learn from the game. Applying this concept, we assume that the SG-variables are mediators between students and facilitation on the one hand and learning and satisfaction on the other hand. Students, by their engagement and functions within teams affect the comprehension and practical relevance of the SG (learning about the game). And by that they can learn as a next step from the game. As it is one of the main responsibilities of game-facilitators to introduce the SG (Kortmann & Peters, 2021) we assume that facilitation helps students to gain a good understanding of the game (learning about the game) so that students can learn from the game.

We therefore introduce six hypotheses:

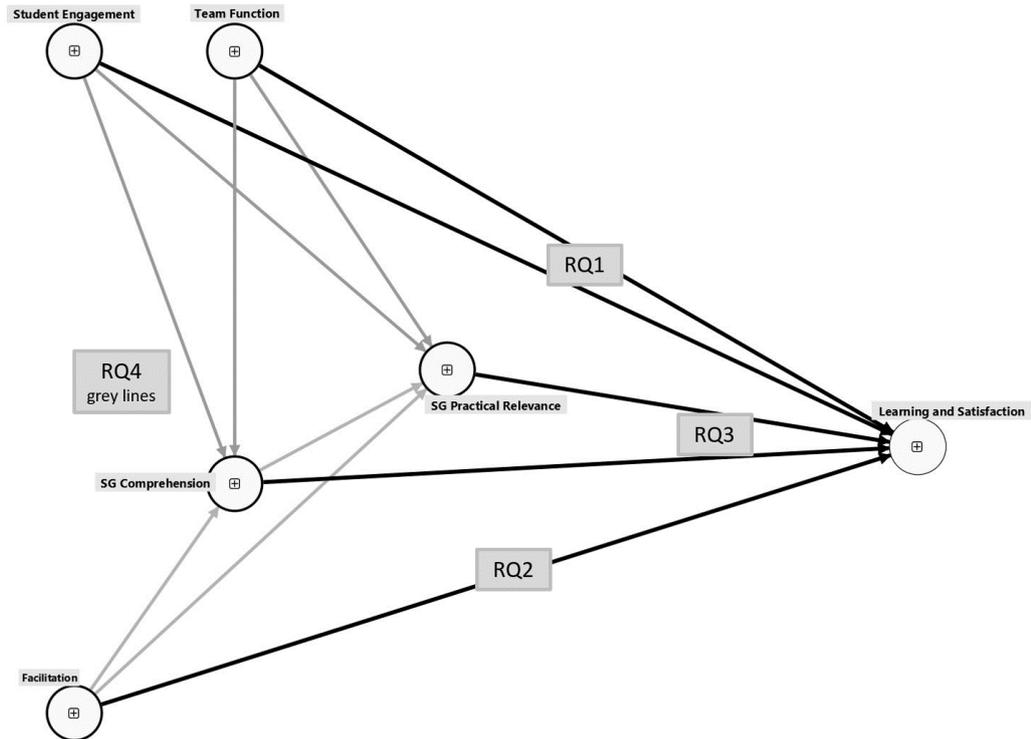
- Student Engagement has a positive effect on SG Comprehension.
- Student Engagement has a positive effect on SG Practical Relevance.
- Team Task Orientation has a positive effect on SG Comprehension.
- Team Task Orientation has a positive effect on SG Practical Relevance.
- Facilitation has a positive effect on SG Comprehension.
- Facilitation has a positive effect on SG Practical Relevance.

In addition to these hypotheses, which describe the concept of learning about the game, we assume a relationship between the two SG variables. To estimate the fidelity or practical relevance of a SG one first must learn and understand the game. Therefore, we assume:

- SG Comprehension has a positive effect on SG Practical Relevance.

Figure 1 illustrates the full model that is tested in this research. On the right we see the dependent variable learning and satisfaction. With student engagement and team function we evaluate RQ1, the impact of students. With the two variables in the middle, the impact of SGs is analyzed (RQ3). With facilitation on the bottom of the figure the impact of facilitation on learning and satisfaction (RQ2) is modeled. In grey lines it also contains the hypotheses that were introduced to analyze RQ4. With the hypotheses pointing from students and facilitation on the SG-variables learning about the game is modelled. With the hypotheses pointing from the SG variables on learning and satisfaction learning from the game is modelled. With this model we can not only estimate the direct effects of students and facilitators on learning but also the indirect effects mediated by the SG variables.

Figure 1: Model: RQs and hypotheses



Data and Methods

Here we introduce the empirical data used for this study and how it was analyzed in a Structural Equation Model using SmartPLS4.

Data Gathering

Data was gathered between May 2021 and April 2023 in bachelor classes at the Faculty of Business and Health at the XXX University (see Table 1). All courses were facilitated at a specialized center for teaching and learning with SGs. Included are 3067 datasets collected in 192 courses that were taught using SGs. All courses were organized as block-seminars meaning that students spend one to three days only playing and debriefing the SG. Lecturers were always present for briefing, debriefing or questions in between. SGs were used with considerably different frequencies (overview on SGs, see appendix). The most frequently used SG is a computer based general management game (34 times) followed by a mainly haptically played SG on change management (20 times). Other games that address very specific topics (like insurance or logistics) were played only once. Researchers did not make choices on SGs included in this study but analyzed what

was played and evaluated during the period of data gathering. Questionnaires were administered digitally towards the end of the courses.

Table 1: Composition of participants and courses

	Options	Number (%)
Gender	Female	1513 (49,3%)
	Male	1395 (45,5%)
	Divers	31 (1.0%)
	No answer	128 (4.2%)
Year of Study	First year (1./2. semester)	1076 (35%)
	Second year (3./4. semester)	640 (21%)
	Third year (5./6. semester)	1245 (41%)
	Not identified	106 (3%)
Area of Study	Business	2694 (88%)
	Health	279 (9%)
	Not identified	94 (3%)
Games	Total	32 (100%)
	Digital games using computer interfaces	19 (59%)
	Haptic games using materials like gameboards/cards/coins	13 (41%)
Courses	Online teaching	85 (44%)
	Onsite teaching	105 (55%)
	Mixture of online and onsite	2 (1%)

Descriptions of Items and Constructs

As mentioned earlier data was gathered with a questionnaire developed for the evaluation of SG-based teaching (F. Trautwein & Alf, 2023). It consists of seven constructs that are formed of 25 scaled items. To investigate the above outlined research questions only six of the seven constructs were used. All questions were answered on a six-point scale, mostly ranging from “do not apply at all” (value = 1) to “fully apply” (value = 6). All items, constructs and indicators on reliability are included in Table 2. Two scales deal with aspects of SGs: how well the game is

understood by the learners (SG-Comprehension) and how well the SG represents relevant real-world issues (SG-Relevance). One construct address teaching specific aspects of facilitation as structure of the seminar or communication between instructors and students. The team-function scale provides information about functions and tasks of students within the team. The engagement scale indicates how intensely students are involved in the game. In the logic of the questionnaire learning and satisfaction is the dependent variable because it represents the goal of teaching. It consists of three items related to learning and three items related to overall satisfaction.

Table 2: Constructs, Items, Reliability (α =Cronbach's alpha, CR=Composite Reliability), Convergence Validity (AVE=Average Variance Extracted) and CFA

Construct	Item	Reliability
SG-Comprehension	S1 The results of the simulation are very comprehensible.	α : 0.830
	S2 The results of the simulation are plausible.	CR: 0.836
	S3 I understood how the simulation works.	AVE: 0.630
SG-Relevance	S4 The simulation has a close practical orientation.	α : 0.854
	S5 The simulation game is a good representation of reality	CR: 0.854 AVE: 0.746
Facilitation	D1 The instructor is very familiar with the simulation.	
	D2 The seminar was well structured.	
	D3 There was sufficient opportunity to consult with the instructor	α : 0.877
	D4 Consultation with the instructor took place in a constructive atmosphere.	CR: 0.877
	D5 Support of the instructor was professionally helpful.	AVE: 0.591
Team-Function	T5 Everybody in the group had a function.	α : 0.818
	T6 There was a clear distribution of tasks in the group.	CR: 0.8153
	T7 I knew what my tasks were.	AVE: 0.599
Student-Engagement	Std2 I actively participated in the decision-making phases (work in small groups).	α : 0.801
	Std3 When others look at my role, they would count me as one of the more active participants	CR: 0.804 AVE: 0.673
Learning-and-Satisfaction	L1 I learned a lot in the simulation game.	
	L2 I learned something meaningful and important in this seminar.	
	L3 My understanding of my field of study has been enhanced by the seminar.	α : 0.939
	G1 All in all, attending the seminar was worth it for me.	CR: 0.939
	G2 I found the simulation game very motivating.	AVE: 0.719
	G3 How satisfied are you with the seminar overall?	

Fit Indicators of Confirmatory Factor Analysis (CFA) for the outlined structure with recommended threshold in brackets:

CFI=0.934 (> 0.9); TLI=0.922 (> 0.9); SRMR=0.0423 (< 0.08); RMSEA=0.0627 (< 0.08)

Partial Least Square Structural Equation Modelling (PLS-SEM)

To analyze the outlined model (Figure 1), we use Structural Equation Modelling (SEM). SEM allows the modelling of networks of causal relationships and therefore can model a whole theory at a time (Lowry & Gaskin, 2014, 125). Under the roof of SEM two families can be distinguished that complement each other: Covariance Based SEM (CB-SEM) and Partial Least Square SEM (PLS-SEM) (Ramli, 2019). CB-SEM follows the idea of factor analysis and is suitable for testing established theory (“hard modelling”) (Ramli, 2019, 302; Weiber & Mühlhaus, 2014, 65) while PLS-SEM follows the idea of regression analysis and can be used to build and explore theory (Hair et al., 2019). All SEMs consist of two models: First, the measuring model or exogenous model that measures the relationship between the observed variables and the constructs. It measures the reliability and validity of the data used for the model. The other is the path model or endogenous model that measures the assumed relationships between the variables. For the interpretation of CB-SEM the overall model fit is emphasized describing how well the outlined theoretical model fits to the structure of the data used (Berning, 2019; Dash & Paul, 2021; Ramli, 2019). PLS-SEM focusses usually on the strength of the assumed relations and the amount of variance in the exogenous constructs that is explained by the endogenous constructs (R^2). In addition, also model fit indicators are available in PLS modelling, but they usually play a minor role in the interpretation compared to CB-SEM (Hair et al., 2019; Ramli, 2019; Ringle et al., 2022). This study uses PLS-SEM mainly for two reasons: First, the evaluation data used is not normally distributed and PLS-SEM is suitable for not normally distributed data (Hair et al., 2019, 5). Second, in the development of the model the authors follow a more exploratory approach (“soft modelling”) modelling many indicators and relationships for which PLS-SEM is the appropriate tool (Hair et al., 2019, 6; Lowry & Gaskin, 2014, 130; Weiber & Mühlhaus, 2014, 72). The model is built and analyzed using SmartPLS 4 (Ringle et al., 2022). It contains only reflective factors (arrows pointing from the latent construct to indicators). Therefore, the consistent PLS-SEM algorithm and the consistent PLS-SEM bootstrapping is used for calculation (Ringle et al., 2022). In case of missing values case wise deletion was used.

For the interpretation of results, we report indicators for PLS-SEM as they are recommended (Hair et al., 2019; Henseler et al., 2016; Lowry & Gaskin, 2014). The relevant indicators and thresholds are outlined in Table 3:

Table 3: Indicators for the interpretation of the model and thresholds

		Indicator	Threshold
Outer Model	Internal Consistence / Reliability	Cronbach's Alpha (α)	>0.7
		Composite Reliability (CR)	>0.7
	Convergent Validity	Average Variance Extracted (AVE)	>0.50
	Discriminant Validity	Heterotrait-Monotrait Ratio (HTMT)	<0.85
Inner Model	Variance Explained	Adjusted R ²	0.75, 0.50 and 0.25 are considered to be substantial, moderate, and weak.
Model Fit		Standardized Root Mean Square Residual (SRMR)	<0.08 indicates a solid fit
		Normed Fit Index (NFI)	>0.9 indicates a solid fit

Outer Model: Reliability and Validity

The internal consistency is an indicator for reliability and answers the question how consistent the single items represent the construct they stand for. The relevant indicators were already presented in Table 2. Both, Cronbach's alpha and Composite Reliability indicate reliable constructs.

The validity of a scale indicates how well the data measure what they are supposed to measure. We report AVE for convergent validity and HTMT for discriminant validity. The AVE describes how much variation in the items is explained by the factor they represent. For all constructs this should be over 50% (> 0.5) which is true for the data used in this study (Table 2). Discriminant validity shows that constructs are distinguished from other constructs. The HTMT indicates that this is true when all ratios are smaller than 0.85, which is the case here (Table 4).

Table 4: HTMT – Discriminant Validity

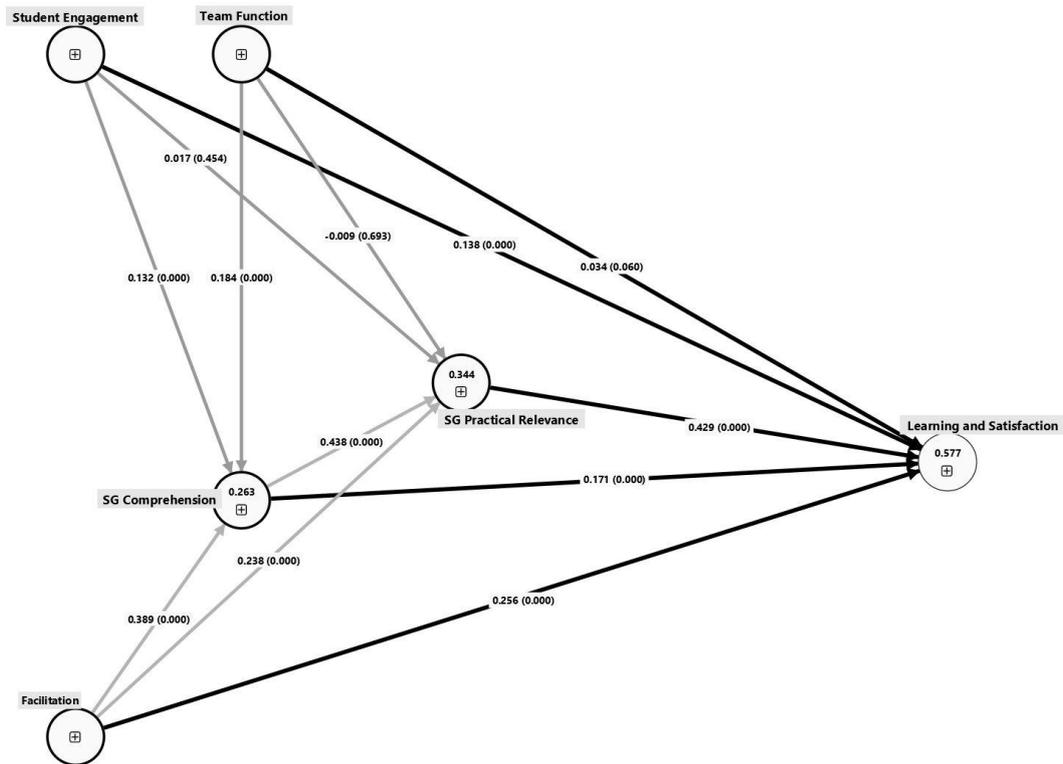
	1	2	3	4	5	6
1. Facilitation						
2. Learning and Satisfaction	0.547					
3. SG Comprehension	0.448	0.568				
4. SG Relevance	0.435	0.666	0.550			
5. Student Engagement	0.214	0.334	0.289	0.188		
6. Team Task	0.146	0.239	0.290	0.157	0.334	

In addition, CFA was conducted using JAMOVI (The jamovi project, 2023) to further validate the outer model. Using various fit-indicators a solid factor structure can be reported (Table 2). Overall the outer model is reliable and valid and can be used for further interpretation.

Results

To give an overview over the results, we first present the overall tested model and then answer the outlined research questions with the relevant hypotheses. Figure 2 and table 5 give an overview over all hypotheses tested with the model.

Figure 2: PLS-SEM: SG-based teaching and learning; path coefficient, p-value and adjusted R²



Model

Looking at the adjusted R², we find lower values for SG-Comprehension (0.236) and SG Practical Relevance (0.344). These lower values are not surprising because in the logic of the model these are actually independent variables. By the interdependent network of hypotheses, they became (also) dependent variables, but it was not the main idea of the model to explain them. An adjusted R² of 0.577 is found for learning and satisfaction which means that the model explains 58% of the variance in the variable learning and satisfaction. Referring to the above introduced thresholds this is a moderate result.

With 0.041 the SRMR indicates a solid model fit while the NFI (0.876) is slightly lower as recommended.

RQ1: How do students affect learning and satisfaction?

Regarding RQ1 we introduced the two hypotheses that both, student engagement and team function would affect learning and satisfaction. But we find that only

student engagement affects learning and satisfaction while team function has no significant impact on learning and satisfaction. The impact of student engagement is highly significant but rather low. The results on RQ1 leave an ambivalent picture that needs to be discussed further.

RQ2: How do facilitators affect learning and satisfaction?

This research question was tested with one hypothesis that is accepted. Facilitation has a highly significant moderate impact on the variable learning and satisfaction.

RQ3: How do simulation games affect learning and satisfaction?

Two hypotheses were used to analyze RQ3. Both hypotheses show a congruent picture. SGs significantly affect learning and satisfaction. The impact on learning and satisfaction from SG comprehension is significant but rather low. The practical relevance of SGs has the highest direct impact (0.429) on learning and satisfaction that is reported in this study. Compared to students and facilitation the SGs' practical relevance is essential for learning and satisfaction.

RQ4: How do Students, Teachers and Simulation Games interact with each other and affect learning and satisfaction?

Under this research question seven hypotheses were tested (grey lines in figure 3). Following the distinction of learning about the game and learning from the game we expected the SG variables to be mediators from students and facilitation on learning and satisfaction.

The two student variables (student engagement and team function) significantly impact SG comprehension (with rather low path coefficients). Engagement and having a function within the team obviously helps students to understand the simulation and to know how to play the game. But student engagement and team function do not impact the practical relevance of the SG. The SG does not appear more realistic and relevant caused by student behavior. The two hypotheses from facilitation on SG comprehension and SG practical relevance are both significant

and demonstrate stronger effects than the student variables. Especially SG comprehension is strongly influenced by facilitation. Finally, we find a significant and comparatively high impact from SG comprehension on SG practical relevance meaning that understanding a SG helps to see it's relation and relevance for reality.

Using SEM allows to estimate indirect effects. These describe effects from one variable to a dependent variable mediated by a third (or fourth) variable (for all indirect effects see appendix). The indirect effect from facilitation on learning and satisfaction (0.242; $p=0.000$) mediated by the two SG variables is almost as high as the direct effect (0.256; $p=0.000$). The indirect effects from student engagement and team function are both highly significant but rather low (<0.01).

Estimating the total effects (direct effect plus indirect effect) facilitation has the highest total impact on satisfaction and learning (direct: 0.256 + indirect: 0.242 = 0.498) that is found in this model.

Table 5: Direct Effects

	Hypotheses	Path coefficient	P value
RQ1	Student Engagement -> Learning and Satisfaction	0.138	0.000
	Team Function -> Learning and Satisfaction	0.034	0.060
RQ2	Facilitation -> Learning and Satisfaction	0.256	0.000
RQ3	SG Comprehension -> Learning and Satisfaction	0.171	0.000
	SG Relevance -> Learning and Satisfaction	0.429	0.000
RQ4	Student Engagement -> SG Comprehension	0.132	0.000
	Student Engagement -> SG Relevance	0.017	0.454
	Team Function -> SG Comprehension	0.184	0.000
	Team Function -> SG Relevance	-0.009	0.693
	Facilitation -> SG Comprehension	0.389	0.000
	Facilitation -> SG Relevance	0.238	0.000
	SG Comprehension -> SG Relevance	0.438	0.000

Discussion

This study addresses four research questions in the context of teaching and learning with SGs in higher education: How do students impact learning and satisfaction? How do facilitators impact learning and satisfaction? How do SGs impact learning and satisfaction? And how do the three independent variables interact with each other to impact learning and satisfaction? Based on data

collected with a questionnaire to evaluate SG-based lecturing data from 192 courses were analyzed to answer the research questions. The exogenous pls model meets the required thresholds for reliability and validity and can be used to analyze the endogenous model. The results briefly described in the last paragraph are now discussed further.

RQ1: How do students affect learning and satisfaction?

As expected, student engagement significantly impacts learning and satisfaction. Compared to the strong relationship between engagement and learning as presented in the literature review the rather low path coefficient (0.138) surprises. However, it should be considered that the variable SG comprehension describes how well students understand the SG and the results given by it. Therefore, this construct has a lot in common with student's mental processes of learning how the SG works. Questioning participants influence on learning and satisfaction we should keep in mind, that student involvement is also represented in SG-comprehension, which will be discussed later. Unexpected is also that team function does not affect learning and satisfaction. According to the data analysed it doesn't matter with regard to learning and satisfaction whether students have the impression that they fulfil a clear task or a function within their team or not. As already mentioned in the literature review team variables were in all studies used as predictors for game-performance. Further research should examine the relationship of game-performance and learning – and especially the impact of team variables on these. It could be that teamwork variables are an important factor for performance but of minor relevance for the individual assessment of learning and satisfaction.

RQ1 can be answered as such: Students do impact learning and satisfaction with their activity, but to a lower extend as expected.

RQ2: How do facilitators affect learning and satisfaction?

Facilitation has a significant moderate direct impact on learning and satisfaction. Measured by the prominent status that the topic of facilitation is given in literature the direct impact of 0.256 seems to be surprisingly small. Under RQ4 die indirect effects of facilitation will be discussed further.

RQ3: How do simulation games affect learning and satisfaction?

The two SG variables have a significant direct impact on learning and satisfaction, but the strength of the impact varies mentionable. While SG comprehension has a small impact of 0.171 the impact of SG practical relevance is with 0.429 the highest direct impact that is found in this study. These findings are congruent with prior research. SG variables like fidelity or reality are in various studies found to be the most important predictors for learning, as outlined in the literature review.

RQ4: How do students, facilitators and simulation games interact with each other and affect learning and satisfaction?

RQ3 revealed that as expected SGs and especially their practical relevance have a very strong impact on learning and satisfaction. Overall, this is a contra intuitive result. SGs are usually very complex and not very self-explanatory. Why would they be so helpful for learning? With RQ4 we therefore evaluate how SGs can exert such influence. Among the seven tested hypotheses under RQ 4 we find two to be not significant. The two student variables (student engagement and team function) have no impact on the practical relevance of the SG, but both impact SG comprehension. Facilitation has a significant impact on SG relevance but a way higher impact on SG comprehension. All impact factors (engagement, team function, facilitation) show the same pattern: They have no or a minor impact on SG practical relevance but a significant or essential impact on SG-comprehension. Through facilitation and their own activity students understand and learn how to play SGs – and thereby SGs can reveal their practical relevance. The impact from SG comprehension on SG relevance is with 0.438 the highest direct effect found in this study. That SGs are found to be the most important predictors for learning requires students to learn about the SGs and it needs especially facilitators that enable students to understand and play them. Only then SGs can unfold their educational potentials. Considering this background, it becomes understandable that overall facilitation has the highest total impact (direct + indirect effect) on satisfaction and learning. Educators always have a direct impact on learning but using complex learning environments like SGs they have also a very relevant

indirect impact. By enabling students to learn about the SG they help them to learn from the SG.

Conclusion

As Hofstede et al. we also were not able to find the “active substance” that is at work when learning with SG, because it probably doesn’t exist. But we found significant impacts from student engagement, from facilitation and from the SG variables on learning and satisfaction. This confirms that teaching and learning are multidimensional constructs that cannot be explained by single indicators (Rindermann, 2009). Comparing the strength of the impact students, facilitators and SGs take on learning and satisfaction it looks at a first glance that especially the practical relevance of the SG is the most important predictor. This result is in line with C. Trautwein (2011), Zeiner-Fink et al. (2023) and Hense et al. (2009), who found that variables related to the SG like realism, fidelity or appropriateness are the best predictors for learning with SGs. Going beyond the direct effects this study also analyses the relationships between the independent variables (the impact of students and facilitation on the SG variables), which most other studies do not. Analysing these relationships reveals the central role of facilitation. Facilitation has a strong impact on how students interpret the SG variables. So, the fidelity/realism/practical relevance might be the most important direct impact factor for learning. But it only becomes such an important factor by facilitators that introduce the SG, guide the process, and debrief the experience. Further research should examine this relationship further. Here it could be distinguished between games that are designed very close to reality and others that remain more abstract. It could be evaluated how this distinction impacts the relationship between facilitation and realism and so the moderation of the SG variables on learning and satisfaction.

The results found in this study are valid for SGs in a broad sense, as it includes evaluation data from over 30 different games. Further research should investigate whether the effects are stable for certain types of SGs. For example, the model could be compared for board games and computer-based games, or rather short vs. longer games.

On a methodological level this study emphasises the potentials of SEM in comparison to regression modelling. Many studies highlight the impact of SG variables on learning when using regression. This is also the case when using SEM. But SEM can also reveal interdependencies between the interdependent variables while this is usually not the case in regression modelling.

Limitations

The data was collected with students of XXX University, Faculty of Business and Health. Therefore, it remains open, whether the results can be transferred to SG lectures in other educational contexts. Although data from 32 different SG are included in the study, the focus lies on SG in the wider area of business. Therefore, future research should examine whether the results are valid for SGs in a wider sense, e.g., political SGs.

A remark must be made on perception and objective measurement of learning. Student's perception is not a perfect view on learning and objective data is preferable (Clayson, 2009). However, the aim of this study was not measuring objective learning outcomes but to evaluate on a broad data basis the correlative structures that enable learning with SG.

References

- Alas, R., Kross, P., Liivat, A., & Saar, J. (2018). Success factors for teams in business game Dynama. *Problems and Perspectives in Management*, 16(1), 309–319. [https://doi.org/10.21511/ppm.16\(1\).2018.30](https://doi.org/10.21511/ppm.16(1).2018.30)
- Alf, T. (2022). Gelingensbedingungen von Planspiellehrveranstaltungen – Ein Systematic Literature Review. *Die Hochschullehre*, 8, 477–480. <https://doi.org/10.3278/HSL2233W>
- Alf, T., & Trautwein, F. (2023). Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre. *Zeitschrift Für Hochschulentwicklung*, 18, 345–363.
- Berning, C. C. (2019). Strukturgleichungsmodelle. In M. Apelt, I. Bode, R. Hasse, U. Meyer, V. V. Grodeck, M. Wilkesmann, & A. Windeler (Eds.), *Springer reference Sozialwissenschaften. Handbuch Organisationssoziologie* (pp. 1–18). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-16937-4_30-2

- Brazhkin, V., & Zimmerman, H. (2019). Students' Perceptions of Learning in an Online Multiround Business Simulation Game: What Can We Learn from Them? *Decision Sciences Journal of Innovative Education*, 17(4), 363–386. <https://doi.org/10.1111/dsj.12189>
- Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn? *Journal of Marketing Education*, 31(1), 16–30. <https://doi.org/10.1177/0273475308324086>
- Corbeil, P., & Laveault, D. (2011). Validity of a Simulation Game as a Method for History Teaching. *Simulation & Gaming*, 42(4), 462–475. <https://doi.org/10.1177/1046878108325451>
- Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, 173, 121092. <https://doi.org/10.1016/j.techfore.2021.121092>
- Dewey, J. (1938). *Experience & Education. The Kappa Delta Pi Lecture Series*. Free Press.
- Dick, G. N., & Akbulut, A. Y. (2020). Innovative Use of the ERPsim Game in a Management Decision Making Class: An Empirical Study. *Journal of Information Technology Education: Research*, 19, 615–637.
- Dube, E. A. (2015). Using Games in Business Education: An Evaluation Experiment Comparing Games to Other Selected Methods in Teaching Sustainable Development Concepts. *International Journal for Business Education*, 155(1). <https://doi.org/10.30707/IJBE155.1.1648133093.381331>
- Duke, R. D., & Geurts, J. (2004). *Policy games for strategic management: Pathways to the Unknown*. Dutch University Press.
- Geithner, S., & Menzel, D. (2016). Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. *Simulation & Gaming*, 47(2), 228–256. <https://doi.org/10.1177/1046878115624312>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hense, J., Kriz, W. C., & Wolfe, J. (2009). Putting Theory-Oriented Evaluation Into Practice. *Simulation & Gaming*, 40(1), 110–133. <https://doi.org/10.1177/1046878107308078>
- Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, 116(1), 2–20. <https://doi.org/10.1108/IMDS-09-2015-0382>
- Herzig, M. (2019). Fame and Fortune: Developing a Simulation Game for the Music Industry Classroom. *Journal of the Scholarship of Teaching and Learning*, 19(5), 105–122. <https://doi.org/10.14434/josotl.v19i5.24276>
- Hofstede, G. J., Caluwé, L. de, & Peters, V. (2010). Why Simulation Games Work – In Search of the Active Substance: A Synthesis. *Simulation & Gaming*, 41(6), 824–843. <https://doi.org/10.1177/1046878110375596>
- Hwang, M. I. (2018). Relationship between Teamwork and Team Performance: Experiences from an ERPsim Competition. *Journal of Information Systems Education*, 29(3), 157–167.
- The jamovi project. (2023). *jamovi* (Version 2.3) [Computer software]. <https://www.jamovi.org>

- Klabbers, J. H. G. (2018). On the Architecture of Game Science. *Simulation & Gaming*, 49(3), 207–245. <https://doi.org/10.1177/1046878118762534>
- Kortmann, R., & Peters, V. (2021). Becoming the Unseen Helmsman - Game facilitator competencies for novice, experienced, and non-game facilitators. *Simulation & Gaming*, 52(3), 255–272. <https://doi.org/10.1177/10468781211020792>
- Kriz, W. C. (2010). A Systemic-Constructivist Approach to the Facilitation and Debriefing of Simulations and Games. *Simulation & Gaming*, 41(5), 663–680. <https://doi.org/10.1177/1046878108319867>
- Kriz, W. C., & Manahl, W. (2016). Understanding and Changing Systems Through Hybrid Simulation Game Design Methods in Educational Contexts. In T. Kaneda, H. Kanegae, Y. Toyoda, & P. Rizzi (Eds.), *Translational Systems Sciences. Simulation and Gaming in the Network Society* (Vol. 9, pp. 79–93). Springer Singapore. https://doi.org/10.1007/978-981-10-0575-6_7
- Lowry, P. B., & Gaskin, J. (2014). Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It. *IEEE Transactions on Professional Communication*, 57(2), 123–146. <https://doi.org/10.1109/TPC.2014.2312452>
- Mandl, H. (2004). On the path to a new learning culture - the contribution of web-based business games. In T. Eberle (Ed.), *Bridging the gap: Transforming knowledge into action through gaming and simulation: Proceedings of the 35th Conference of the International Simulation and Gaming Association, Munich, 2004* (pp. 163–174).
- Mayer, I., Warmelink, H., & Bekebrede, G. (2013). Learning in a Game-Based Virtual Environment: A Comparative Evaluation in Higher Education. *European Journal of Engineering Education*, 38(1), 85-106. <https://doi.org/10.1080/03043797.2012.742872>
- Pasin, F., & Giroux, H. (2011). The Impact of a Simulation Game on Operations Management Education. *Computers & Education*, 57(1), 1240–1254. <https://doi.org/10.1016/j.compedu.2010.12.006>
- 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>
- Ramli, M. Z. (2019). A Review of Structural Equation Model for Construction Delay Study. *International Journal of Engineering & Technology*, 7(4), 299–306.
- Richards, B., & Camuso, L. (2015). Cultural Capital in the Classroom: The Significance of Debriefing as a Pedagogical Tool in Simulation-Based Learning. *International Journal of Teaching and Learning in Higher Education*, 27(1), 94–103.
- Rindermann, H. (2009). *Lehrevaluation: Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen mit einem Beitrag zur Evaluation computerbasierter Unterrichts* [2., leicht korrigierte Aufl.]. *Psychologie: Vol. 42*. Verl. Empirische Pädagogik.
- Ringle, C. M., Wende, S., & Becker, J.-M. (2022). *SmartPLS 4*. <https://www.smartpls.com/>

- Tao, Y.-H., Yeh, C. R., & Hung, K. C. (2012). Effects of the Heterogeneity of Game Complexity and User Population in Learning Performance of Business Simulation Games. *Computers & Education*, 59(4), 1350–1360. <https://doi.org/10.1016/j.compedu.2012.06.003>
- Trautwein, C. (2011). *Unternehmensplanspiele im industriebetrieblichen Hochschulstudium: Analyse von Kompetenzerwerb, Motivation und Zufriedenheit am Beispiel des Unternehmensplanspiels TOPSIM - General Management II*. Zugl.: Hohenheim, Univ., Diss., 2010 (1. Aufl.). Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden. <https://doi.org/10.1007/978-3-8349-6955-2>
- Trautwein, F., & Alf, T. (2023). Theory-Based Development of an Inventory for the Evaluation of Simulation Game Lectures. In C. Harteveld, S. Sutherland, G. Troiano, H. Lukosch, & S. Meijer (Eds.), *Simulation and Gaming for Social Impact* (Vol. 13622, pp. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-031-37171-4_1
- van Blisen, A., Bekebrede, G., & Mayer, I. (2010). Understanding Complex Adaptive Systems by Playing Games. *Informatics in Education*, 9(1), 1–18. <https://files.eric.ed.gov/fulltext/EJ1064272.pdf>
- van den Bossche, P., Gijssels, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team Learning: Building Shared Mental Models. *Instructional Science*, 39(3), 283–301.
- Wijse-van Heeswijk, M. de. (2023). Facilitation Interventions to Increase Learning Effectiveness in Game Simulations. A Generic Approach of Facilitation Applicable to a Broad Variety of Simulation Games. In M. L. Angelini & R. Muñiz (Eds.), *Springer Texts in Education. Simulation for Participatory Education* (pp. 53–85). Springer International Publishing. https://doi.org/10.1007/978-3-031-21011-2_4
- Weiber, R., & Mülhau, D. (2014). *Strukturgleichungsmodellierung*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-35012-2>
- Wijse-van Heeswijk, M. de. (2023). Facilitation Interventions to Increase Learning Effectiveness in Game Simulations. A Generic Approach of Facilitation Applicable to a Broad Variety of Simulation Games. In M. L. Angelini & R. Muñiz (Eds.), *Springer Texts in Education. Simulation for Participatory Education* (pp. 53–85). Springer International Publishing. https://doi.org/10.1007/978-3-031-21011-2_4
- Zeiner-Fink, S., Bullinger, A. C., & Geithner, S. (2023). Learning Effects and Acceptance in Business Games: A Systematic Literature Review. In C. Harteveld, S. Sutherland, G. Troiano, H. Lukosch, & S. Meijer (Eds.), *Simulation and Gaming for Social Impact* (Vol. 13622, pp. 36–51). Springer International Publishing. https://doi.org/10.1007/978-3-031-37171-4_3

Appendix

Simulation Games

Simulation Game	Main Topic	Number of cases with this game	Percent %
TOPSIM General Management	General Management, Business Strategy	621	20.25
riva SysTeamsChange	Change Management	372	12.13
TOPSIM Easy Management	General Management, Business Strategy	230	7.50
BTI Factory	Accounting, Business	222	7.24
AMGD	Project Management, Scrum	195	6.36
BTI Global Strategy	Accounting, Business Strategy	179	5.84
coludo-Industry	Accounting, Business Strategy	153	4.99
Fishing Game	Sustainability, Social Responsibility, Leadership	109	3.55
CAPSIM-Core	General Management, Business Strategy	108	3.52
TOPSIM Logistics	Management, Logistics	101	3.29
BTI Hospital Akut	Accounting, Health System	92	3.00
Ökonomikus Insurance	Business Strategy, Insurance	83	2.71
BTI Store Manager	Accounting, Business Strategy, Retail	77	2.51
TOPSIM Universal Banking	Business Strategy, Banking	71	2.31
TOPSIM Insurance	Business Strategy, Insurance	58	1.89
TOPSIM Hospital Management	Business Strategy, Health System	55	1.79
Grid	Organizational Development, Collaboration, Leadership	36	1.17
Serious Games.Studio - Tool Instrument GmbH	Business Strategy, Accounting	31	1.01
TOPSIM Applied Economics	Macro Economics	30	0.98
BTI Logistic	Accounting, Logistics	29	0.95
Fountains	Sustainability, Social Responsibility, Leadership	27	0.88
HR-Management	Business Strategy, HR-Management	26	0.85
TOPSIM Immo Makler	Business Strategy, Real Estate	24	0.78
TOPSIM BD Fitness	Business Strategy, Fitness	22	0.72
Inchaing - The Fresh Connection	Business Strategy	19	0.62
Beergame	Logistics	18	0.59
Mission Team Development	Leadership, Teambuilding, Cooperation	16	0.52
TOPSIM Manufacturing Management	Business Strategy, Manufacturing	15	0.49
TOPSIM Scale Up	Business Strategy, Disruptive Markets	15	0.49

IT-Service-Management	Business Strategy, IT-Service	14	0.46
TOPSIM Going Global	Business Strategy, New Market, global Expansion	10	0.33
ProCom Medieval Market	Organizational Development in Hospitals	9	0.29
Total		3067	100

Indirect Effects

Specific Indirect Effects

	Path coefficient	P-values
Facilitation -> SG-Comprehension -> SG-Relevance	0.170	0.000
Student-Engagement -> SG-Relevance -> Learning-and-Satisfaction	0.007	0.455
Student-Engagement -> SG-Comprehension -> Learning-and-Satisfaction	0.023	0.000
Team-Task -> SG-Relevance -> Learning-and-Satisfaction	-0.004	0.693
Team-Task -> SG-Comprehension -> Learning-and-Satisfaction	0.031	0.000
Student-Engagement -> SG-Comprehension -> SG-Relevance	0.058	0.000
Team-Task -> SG-Comprehension -> SG-Relevance	0.080	0.000
Team-Task -> SG-Comprehension -> SG-Relevance -> Learning-and-Satisfaction	0.034	0.000
Facilitation -> SG-Comprehension -> SG-Relevance -> Learning-and-Satisfaction	0.073	0.000
Student-Engagement -> SG-Comprehension -> SG-Relevance -> Learning-and-Satisfaction	0.025	0.000
Facilitation -> SG-Relevance -> Learning-and-Satisfaction	0.102	0.000
Facilitation -> SG-Comprehension -> Learning-and-Satisfaction	0.066	0.000
SG-Comprehension -> SG-Relevance -> Learning-and-Satisfaction	0.188	0.000

Total Indirect Effects

	Path coefficient	P-values
Facilitation -> Learning-and-Satisfaction	0.242	0.000
Facilitation -> SG-Relevance	0.170	0.000
SG-Comprehension -> Learning-and-Satisfaction	0.188	0.000

Student-Engagement -> Learning-and-Satisfaction	0.055	0.000
Student-Engagement -> SG-Relevance	0.058	0.000
Team-Task -> Learning-and-Satisfaction	0.062	0.000
Team-Task -> SG-Relevance	0.080	0.000

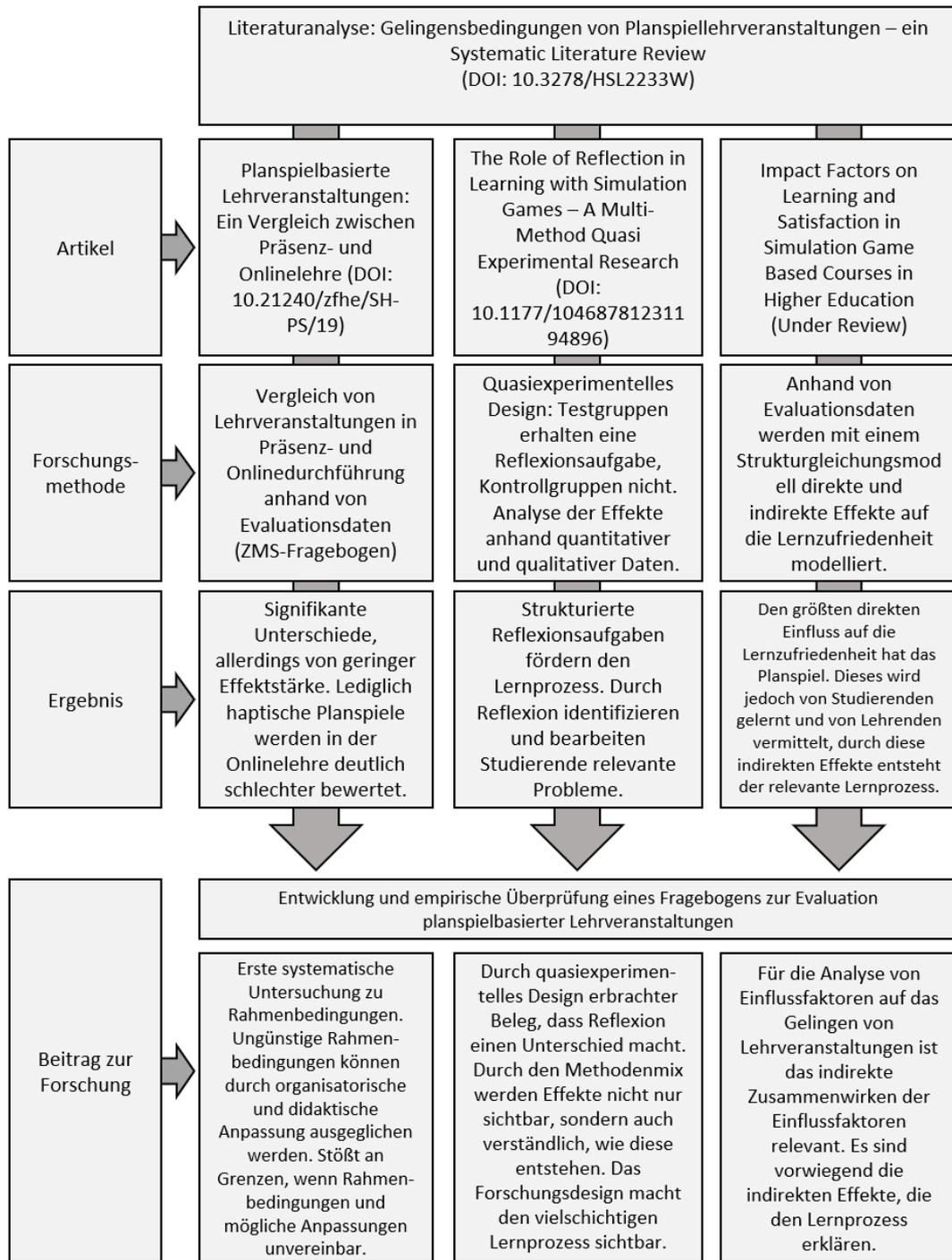
Total Effects (direct effect + total indirect effect)

	Path coefficient	P values
Facilitation -> Learning-and-Satisfaction	0.498	0.000
Facilitation -> SG-Comprehension	0.389	0.000
Facilitation -> SG-Relevance	0.409	0.000
SG-Comprehension -> Learning-and-Satisfaction	0.358	0.000
SG-Comprehension -> SG-Relevance	0.438	0.000
SG-Relevance -> Learning-and-Satisfaction	0.429	0.000
Student-Engagement -> Learning-and-Satisfaction	0.192	0.000
Student-Engagement -> SG-Comprehension	0.132	0.000
Student-Engagement -> SG-Relevance	0.075	0.001
Team-Task -> Learning-and-Satisfaction	0.096	0.000
Team-Task -> SG-Comprehension	0.184	0.000
Team-Task -> SG-Relevance	0.072	0.002

6 Diskussion der Ergebnisse und Beitrag zur Forschung

Nach der einzelnen Betrachtung der Literaturstudie und den daraus abgeleiteten empirischen Studien stellt sich nun die Frage, welche allgemeineren Schlüsse sich hieraus ziehen lassen. Welche Beiträge leisten die Artikel zur Forschung und welche übergreifenden Erkenntnisse lassen sich daraus ziehen? Abbildung 6 gibt hierzu einen Überblick, während die Punkte im Folgenden weiter ausgeführt werden.

Abbildung 6: Beitrag zur Forschung



6.1 Gelingensbedingungen Planspielbasierter Lehrveranstaltungen

Während der entwickelte und in Kapitel 4 vorgestellte Fragebogen die empirische Grundlage dieser Arbeit bildet, kann die systematische Literaturlanalyse als inhaltliche Grundlage dieser Arbeit angesehen werden. Der Artikel analysiert ein breites Spektrum an Forschungsarbeiten rund um Planspiele/Simulation Games und fragt nach Gelingensbedingungen im Hinblick auf Rahmenbedingungen, Lehrende, Planspiele und Studierende. Die Literaturlstudie fördert ein ganzes Bündel an Erkenntnissen zutage. Planspiele werden in sehr vielen unterschiedlichen Fachrichtungen (und mit sehr unterschiedlichen Lernzielen) entwickelt, eingesetzt und beforscht. Zentrale Erkenntnisse bezüglich der Gelingensbedingungen betreffen die Anpassbarkeit von Planspielen an die Zielgruppe, die Authentizität von Planspielen, die Rolle von Lehrenden sowie das Engagement von Studierenden. Fast alle analysierten Studien kommen zu einer positiven Einschätzung hinsichtlich des Lernens mit Planspielen. Der Artikel diskutiert zudem, dass der englische Begriff Simulation Game wesentlich uneinheitlicher gebraucht wird als das deutschsprachige Pendant Planspiel.

Die Literaturlstudie weist in ihrem Fazit insbesondere auf drei Lücken in der bisherigen Forschung hin: Die Rolle von Rahmenbedingungen beim Einsatz von Planspielen wird bislang nicht thematisiert. Die Vielschichtigkeit des Lernens mit/in Planspielen wird aufgrund zu einfacher Forschungsdesigns nur unzureichend abgebildet. Planspielveranstaltungen sollten als multifaktorielle Konstrukte wahrgenommen und untersucht werden. Die empirischen Studien adressieren jeweils eines dieser Themen, insofern dient die Literaturlstudie als inhaltliche Grundlage dieser Arbeit.

6.2 Rahmenbedingungen: Ein Vergleich zwischen Online- und Präsenzlehre

Mit dem Artikel *Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre* erfolgte erstmals eine strukturierte Untersuchung zum Einfluss von Rahmenbedingungen auf die Lehre mit Planspielen. Es wurde deutlich, dass die hier untersuchte Rahmenbedingung in vielerlei Hinsicht Einfluss auf den Lehr-Lernprozess nimmt. Das Dozierendenverhalten wird überraschenderweise in Onlineveranstaltungen etwas besser eingeschätzt als in

Präsenzveranstaltungen. Studierende bewerten ihr Engagement in Onlineveranstaltungen etwas höher, sehen aber Nachteile für die Teamarbeit (Forschungsfrage 1). Für die Bewertung der Planspiele (Forschungsfrage 2) zeigt sich, dass für eine sinnvolle Interpretation der Ergebnisse zwischen computergestützten und haptischen Planspielen unterschieden werden muss. Computergestützte Planspiele werden in der Onlinelehre etwas verständlicher und relevanter für die Praxis eingeschätzt, während haptische Planspiele in der Präsenzlehre als verständlicher und relevanter für die Praxis eingeschätzt werden. Im Hinblick auf Lernen und Zufriedenheit (Forschungsfrage 3) zeigt sich, dass die Lehrveranstaltungen sowohl in der Präsenz- als auch in der Onlinelehre ähnlich bewertet werden, mit Ausnahme der ungünstigen Konstellation von haptischen Planspielen in der Onlinelehre. Angesichts des massiven Einschnitts fallen die Effekte jedoch allenfalls moderat aus. Ungünstige Rahmenbedingungen nehmen demnach nicht zwangsläufig massiv negativen Einfluss auf den Lehr-Lernprozess. Durch didaktische und organisatorische Anpassungen können ungünstige Rahmenbedingungen ausgeglichen werden. Jedoch werden durch die Studie auch Grenzen der didaktischen Anpassung sichtbar. Rahmenbedingungen und didaktische Möglichkeiten können sich beinahe unvereinbar gegenüberstehen. Das Beispiel haptische Planspiele in der Distanzlehre macht dies deutlich. Mit dem Beispiel Präsenz/Online wurde lediglich eine von vielen möglichen Rahmenbedingungen auf die Lehre untersucht. Dieses Feld sollte anhand anderer Beispiele weiter untersucht werden.

Der Artikel macht auch deutlich, dass Online- oder Distanzlehre ein Überbegriff für ein sehr heterogenes Feld ist. Um weitere Differenzierung zu schaffen, greift der Artikel daher auf Entner et al. zurück, die Lehrveranstaltungen nicht nur nach dem Ort (virtuell oder physisch) sondern auch in zeitlicher Hinsicht (synchron oder asynchron) gliedern (2021). Da in die empirischen Daten nur Fälle von synchroner Onlinelehre und synchroner Präsenzlehre eingegangen sind, können im engeren Sinne auch nur hierzu Aussagen getroffen werden. Dennoch ist auffällig, dass einige der in der Literatur diskutierten Probleme der Distanzlehre vorwiegend in asynchronen Veranstaltungen relevant sein dürften. Das Problem der mangelnden persönlichen Kontakte zu Peers und Lehrenden (Xu & Xu, 2020, S. 382) scheint unseren Daten zufolge in der synchronen Distanzlehre mittels Videokonferenz nicht relevant zu sein. Wenngleich weit weniger als während der pandemiebedingten Onlinelehre, bleiben Formate des Distanzlernens sicherlich

auch an Präsenzhochschulen in der ein oder anderen Weise erhalten. Um diese weiter zu erforschen und diskutieren zu können ist es notwendig differenziert zu beschreiben, was eigentlich untersucht wird (synchron/asynchron, virtuell/physisch). Ohne weitere Erklärung von Online- oder Distanzlehre zu sprechen, wird der Vielfalt der möglichen Formate mit ihren Vor- und Nachteilen nicht gerecht.

Die Ergebnisse weisen zudem darauf hin, dass Lehrveranstaltungsevaluation und Lehrevaluation/Studienevaluation (siehe Kapitel 2.4) tatsächlich zwei unterschiedliche Evaluationsgegenstände sind. Während die planspielbasierten Lehrveranstaltungen während der Coronapandemie am ZMS von Studierenden kaum schlechter (in manchen Bereichen sogar minimal besser) bewertet wurden, kann angenommen werden, dass die Bewertung der allgemeinen Studienbedingungen, insgesamt differenzierter ausfällt. So dachten in Deutschland während der Coronapandemie insbesondere benachteiligte Gruppen häufiger über einen Studienabbruch nach (Koopmann et al., 2024) als vor der Pandemie. In einer Interviewstudie unter dual Studierenden wird zudem deutlich, dass sich das Lernen während der Coronapandemie deutlich verändert hat und das Selbststudium einen sehr viel höheren Stellenwert einnimmt (Hettler et al., 2022).

6.3 Strukturierte Reflexionsaufgaben

Während die Rolle der Reflexion bzw. des Debriefings beim Lernen mit Planspielen in eher konzeptionellen Artikeln häufig betont wird (Peters & Vissers, 2004; Schwägele et al., 2021; Wijse-van Heeswijk, 2023), wurde sie bislang kaum empirisch untersucht. Der Artikel *The Role of Reflection in Learning with Simulation Games – A Multi-Method Quasi Experimental Research* untersucht die Wirksamkeit einer Reflexionsaufgabe empirisch und greift dafür auf ein quasiexperimentelles Zweigruppendedesign zurück. Anhand des Designs können Unterschiede zwischen Test- und Kontrollgruppe auf die Intervention (Reflexionsaufgabe) zurückgeführt werden. Als Ergebnis zeigt sich, dass die Rolle von Reflexion und Debriefing beim Lernen mit Planspielen zurecht stark betont wird. Die Testgruppe schätzt ihre Lernergebnisse besser ein, hat ein besseres Verständnis des Planspiels, arbeitet besser im Team und ist aktiver. Zudem erzielt die Testgruppe auch bessere objektive Spielergebnisse. Die qualitativen Daten zeigen, warum diese Ergebnisse plausibel sind. Die Testgruppen identifizierten für

die Simulation relevante Probleme und konnten diese im Verlauf des Spiels häufig erfolgreich bearbeiten. Es ist ein wesentlicher Beitrag dieses Artikels, das Thema Reflexion/Debriefing nicht nur konzeptionell zu fassen, sondern anhand empirischer Daten zu zeigen, dass strukturierte Reflexionsfragen einen wesentlichen Unterschied machen können. Durch das komplexe Forschungsdesign (Test- und Kontrollgruppe, mehrfache quantitative Datenerhebung mit verschiedenen Instrumenten, Einbezug von Spielergebnissen und qualitativen Daten zu verschiedenen Zeitpunkten) ermöglicht die Studie einen Blick auf den Vielschichtigen Lernprozess. Insbesondere durch die qualitative Analyse der Reflexionsprotokolle wird verständlich, wie die Unterschiede zwischen den Test- und Kontrollgruppen zustande kommen und erklärt werden können: Während in früheren Runden relevante Probleme identifiziert werden, können sie in späteren Runden erfolgreich bearbeitet werden. Relevant für die Unterschiede zwischen den Gruppen ist demnach insbesondere die wiederkehrende Möglichkeit von Reflexion und veränderter Handlung.

6.4 Einflussfaktoren auf Lernen und Zufriedenheit

Lehr-Lernveranstaltungen sind ein multifaktorielles Konstrukt (Rindermann, 2009). Mehrere Dimensionen nehmen Einfluss auf den Lehr- bzw. Lernerfolg. Werden in der Lehre erfahrungsorientierte Lernumgebungen eingesetzt, so trifft dies in besonderem Maße zu, da neben den üblichen Dimensionen (z.B. Strukturierung, Kommunikation, Praxisbeispiele) weitere Dimensionen hinzukommen (z.B. Qualität der Lernumgebung, Teamarbeit). Eine beachtenswerte Erkenntnis aus dem Paper Impact Factors on Learning and Satisfaction in Simulation Game-based Courses in Higher Education ist, dass die multiplen Faktoren nicht (nur) isoliert voneinander Einfluss auf den Lehrerfolg (hier gemessen anhand der Lernzufriedenheit) nehmen, sondern auch in ihrem indirekten Zusammenwirken. Auf den ersten Blick (direkte Effekte) scheinen die Planspielvariablen den stärksten Einfluss auf den Lernerfolg zu nehmen. Auch systematische Literaturstudien (Susann Zeiner-Fink et al., 2023) und Studien mit linearen Regressionsmodellen (Hense et al., 2009; C. Trautwein, 2011) kommen zu diesem Ergebnis. Man könnte die Schlussfolgerung ziehen, Lehrende seien in planspielbasierten Lehrveranstaltungen beinahe überflüssig, da die wesentliche didaktische „Arbeit“ vom Planspiel geleistet wird. Die indirekten Effekte zeigen jedoch, dass das Planspiel seine Bedeutung für den Lehrerfolg durch die

vermittelnde Tätigkeit der Lehrenden gewinnt und nur so zum Lehrerfolg beitragen kann. Es braucht Lehrende, damit Studierende das Spiel verstehen können (learning about the game), um in einem zweiten Schritt vom Spiel lernen zu können (learning from the game).

Auf methodischer Ebene verdeutlicht der Artikel die Relevanz von Strukturgleichungsmodellen in der theoriegeleiteten Evaluation. In Regressionsanalysen bleiben die Beziehungen zwischen den unabhängigen Variablen zumeist unberücksichtigt. In der Regel werden nur die Einflüsse der einzelnen Regressoren auf die abhängige Variable thematisiert, nicht aber deren interdependentes Zusammenwirken¹⁷. Indirekte Effekte (learning about the game and learning from the game) bleiben somit verborgen. Strukturgleichungsmodelle hingegen können Beziehungen zwischen unabhängigen Variablen modellieren und deren Einfluss auf die (eigentliche) abhängige Variable. Sie sind daher geeignet, um eine „Theorie des Programms“ zu formulieren und mit hypothesentestenden Verfahren zu überprüfen (Pohlenz, 2008b, S. 72). Ein Blick in die Black Box kann so gelingen. Mit Strukturgleichungsmodellen können nicht nur multikausale Einflussfaktoren abgebildet werden, sondern auch deren interdependentes Zusammenwirken.

6.5 Fragebogenentwicklung und Ansätze zur Weiterentwicklung

Als Vorbereitung für die empirischen Erhebungen im Rahmen dieser Dissertation wurde ein Fragebogen zur Evaluation planspielbasierter Lehrveranstaltungen entwickelt. Dieser wurde am ZMS intensiv eingesetzt und anhand der gewonnenen Daten auf seine Reliabilität und Validität geprüft. Somit steht zum ersten Mal ein empirisch überprüftes Instrument zur Evaluation planspielbasierter Lehrveranstaltungen zur Verfügung, das auch von anderen Forschenden eingesetzt werden kann. In größerem Umfang wurde der Fragebogen bislang nur am ZMS eingesetzt. Wünschenswert für die weitere Validierung sind auch Erhebungen an anderen Hochschularten sowie mit Planspielen, die andere inhaltliche Schwerpunkte setzen, etwa aus dem Bereich der Politischen Bildung.

¹⁷ Mithilfe des Makros PROCESS (<https://processmacro.org/index.html>), das für viele Statistikprogramme verfügbar ist, wäre dies zwar möglich, jedoch ist mir aus dem Kontext Planspiele und Lernen keine Studie bekannt, die eine solche Analyse durchführt.

Das theoretische Modell des Fragebogens beschreibt Lernen und Zufriedenheit als abhängige Variable. Mit der Studie *Impact Factors on Learning and Satisfaction in Simulation Game-based Courses in Higher Education* wurde unter anderem untersucht wie viel Varianz in Lernen und Zufriedenheit die Einflussfaktoren aufklären können. Knapp 60% aufgeklärte Varianz ist für eine sozialwissenschaftliche Fragestellung ein akzeptabler und zu erwartender Wert (Hair et al., 2019, S. 11). Dennoch bleibt mehr als Drittel der Varianz in Lernen und Zufriedenheit unaufgeklärt. Weitere Einflussfaktoren könnten evtl. zu einer höheren Varianzaufklärung beitragen. Mit Blick auf die empirischen Ergebnisse könnten zur Identifikation weiterer Einflussfaktoren zwei Ansätze verfolgt werden. Die Variable Funktion innerhalb des Teams trägt nur minimal zur Varianzaufklärung in der Variable Lernen und Zufriedenheit bei, während andere Studien gerade in der Teamarbeit einen wichtigen Einfluss erkennen (Treen et al., 2016). Anhand weiterer Literaturstudien oder auch anhand induktiver qualitativer Methoden könnte der Einfluss der Teamarbeit weiter untersucht werden. Evtl. finden sich hier übersehene Aspekte von Teamarbeit, die in höherem Maße zur Varianzaufklärung beitragen können. Einen starken Beitrag zur Varianzaufklärung leisten die Variablen zum Planspiel. Auch hier könnten weitere Variablen getestet werden. Der Fragebogen MEEGA+ erhebt beispielsweise ästhetische Merkmale des Spiels, die im ZMS-Fragebogen bislang unberücksichtigt bleiben. Eher technisch ausgerichtete Studien erheben die Usability von Lernspielen (Magylaité et al., 2022). Auch diese Aspekte könnten sowohl auf die Variablen Verständnis des Planspiels als auch auf Lernen und Zufriedenheit Einfluss nehmen – und so zur Varianzaufklärung beitragen.

Durch die zwischenzeitliche Erweiterung des Planspielportfolios am ZMS fällt auf, dass der Fragebogen an einigen Stellen Implikationen enthält, die nicht für alle Planspiele zutreffend sind. Beispielsweise geht er davon aus, dass Planspiele grundsätzlich in Teams gespielt werden. Ein neueres Planspiel am ZMS (Krankenversicherungsplanspiel, hierzu sind noch keine Daten in dieser Arbeit enthalten) wird jedoch nicht in Teams gespielt. Alle Studierenden nehmen als unabhängige Spieler teil. Hierdurch ergibt sich Bedarf an Anpassungen bei einigen nominalen Fragen (z.B. Größe des Teams) und insbesondere bei den Fragen zur Zusammenarbeit im Team, die hier obsolet werden. Durch Konsultationen mit anderen Forschenden fällt zudem auf, dass der Fragebogen Implikationen enthält, die nur am ZMS zutreffend sind. Aufgrund der institutionellen Zuständigkeit des

ZMS für das Thema Planspiele sind diese an der DHBW Stuttgart aus dem normalen Lehrbetrieb „ausgegliedert“. Planspiele finden als gesonderte Blockveranstaltungen am ZMS statt. Sie sind daher in den Studienplänen der Studierenden extra ausgewiesene Lehrveranstaltungen. An den meisten Hochschulen dürfte dies jedoch anders sein. Dort sind Planspiele vermutlich zumeist in „normale“ Lehrveranstaltungen integriert und somit nur Teil einer Lehrveranstaltung, in der auch andere Methoden zur Anwendung kommen. Auch hier ist eine Anpassung des Fragebogens notwendig, soll er auch für Lehrveranstaltungen Anwendung finden, in denen das Planspiel eher Teil einer größeren Einheit ist.

Wie in Kapitel 3.3 geschildert, bieten Planspiele mehrdimensionale, tiefe und systemische Lernmöglichkeiten und es wäre wünschenswert, diese Breite des Lernens erheben und evaluieren zu können. Mit der Umsetzung des Vorhabens einen kurzen Evaluationsbogen zu entwickeln, der für eine Vielzahl von Planspielen passend ist, geht jedoch einher, dass methodische und inhaltliche Abstriche gemacht werden müssen. So wäre es beispielsweise wünschenswert mit Pre-/Post Fragen zu arbeiten und insbesondere den Lernfortschritt differenzierter zu Erheben. Hierfür könnten allgemeine Kompetenzfragebögen verwendet werden (Braun et al., 2008), aber auch sehr planspielspezifische Erhebungsinstrumente. Als Standard der Evaluation am ZMS wird es jedoch aus zeitlichen und ökonomischen Gründen nicht möglich sein, regelmäßig komplexe Erhebungen durchzuführen. Wie bereits in der Studie *The Role of Reflection in Learning with Simulation Games* geschehen, können jedoch für kleinere Stichproben differenziertere Erhebungsdesigns an die Standardevaluation angedockt werden. Ein besonderer Vorteil hierbei ist, dass durch die Standardevaluation ersichtlich wird, ob die kleinere Stichprobe repräsentativ für einen großen Teil der Planspiellehrveranstaltungen steht, oder ob sich in der Stichprobe besondere Effekte ergeben haben.

7 Exkurs: Kritische Reflexion planspiel- und erfahrungsorientierter Lehre

Zum Abschluss dieser Arbeit möchte ich eine Beobachtung aufgreifen, die ich im Laufe der theoretischen und praktischen Beschäftigung mit Planspielen, und auch

darüber hinaus mit erfahrungsorientierten Lehr-Lernmethoden allgemein, gemacht habe. Stichworte wie Erfahrungsorientierung, Game-based-learning oder Gamification sind zumeist positiv konnotiert und eher selten Gegenstand einer kritischen Auseinandersetzung. Beinahe selbstverständlich werden in der gegenwärtigen Literatur erfahrungsorientierte oder spielbasierte Lehr-Lernprozesse als wünschenswert dargestellt. Von den 58 in der Literaturstudie untersuchten Studien führen beispielsweise nur zwei auch eher negative Konsequenzen für Teilnehmende auf, und dies eher am Rande. In systematischer Weise werden Planspiele (soweit mir bekannt) nicht auf (womöglich unbeabsichtigte) negative Konsequenzen hin untersucht. Beispielsweise fragen mehrere Studien, inwiefern Flow-Erlebnisse in Planspielen auftreten (Matute-Vallejo & Melero-Polo, 2019) – keine Studie fragt jedoch systematisch nach dem Auftreten negativer Emotionen. Daher möchte ich mit diesem Exkurs drei Punkte aufgreifen, die in der gegenwärtigen Diskussion nicht oder kaum Beachtung finden: Bildung als Zumutung, Planspiele als verzerrte Modelle der Realität und die Art und Weise, in der Planspiele reflektiert werden.

Zunächst verlangen erfahrungsorientierte und spielerische Lehr-Lernmethoden von Studierenden ein intensiveres Engagement als „normale“ Lehrmethoden das tun. Planspiele verlangen von Studierenden in eine Rolle zu schlüpfen und diese auszuüben. Erfahrungsorientierte Methoden verlangen nicht nur kognitive Mitarbeit, wie klassische Seminare und Vorlesungen das tun, sie verlangen zudem soziales und auch emotionales Engagement. Anhand von zwei Beispielen kann dies illustriert werden: Insbesondere politische Planspiele machen es erforderlich, dass Teilnehmende Rollen übernehmen, die sie womöglich nicht übernehmen wollen. Sie müssen beispielsweise im Sinne einer Partei zu argumentieren, deren Programm sie im echten Leben niemals vertreten wollten. Als zweites Beispiel dienen Type 2-Games (Kategorisierung nach Klabbers, siehe Kapitel 3.4). Sie verwickeln Spielende absichtlich in konflikthafte Situationen und provozieren möglicherweise unangenehme und hitzige Debatten über Ressourcen und deren gerechte Aufteilung. In gewissem Maße leben erfahrungsorientierte Lehr-Lernmethoden davon, dass Lernende Erfahrungen machen, die sie sich nicht selbst ausgesucht haben. Sie werden in Dilemmata und Entscheidungssituationen gebracht, denen sie sich eigentlich nicht stellen wollten. Bildung wird in der Erziehungswissenschaft auch als **Zumutung** diskutiert (Ehrenspeck et al., 2008). Vielleicht gilt dies in besonderem Maße für erfahrungsorientierte Lehr-

Lernmethoden: Lernenden werden Aktivitäten, Rollen, Szenarien und Gruppenarbeiten zugemutet. Mit dem am ZMS entwickelten Fragebogen wird eine ex-post-ante-Frage gestellt. Die Studierenden bewerten am Ende der Veranstaltung, wie interessiert sie vor der Veranstaltung an dem Planspiel waren. Von allen Items des Fragebogens wird dieses Item mit einem Mittelwert von 3,7 (sechsstufige Skala) am negativsten bewertet. Es könnte sein, dass Studierende vor der Veranstaltung genau diese Zumutung befürchten. Dass das Interesse zu Beginn die negativste Bewertung erfährt sagt zugleich, dass die Gesamtbeurteilung positiver ausfällt, sich die Befürchtung in der Regel also nicht bestätigt. Dennoch spricht das eher verhaltene Interesse zu Beginn für eine gewisse Skepsis der Studierenden, sich auf das eher ungewisse (und vielleicht unbequeme) Lehr-Lernformat einzulassen. Und nicht immer gelingt es, diese Skepsis aufzulösen. In einer reflexiven Auseinandersetzung mit einer eigenen Lehrveranstaltung beschreibt Elmes (2019) wie er eine Gruppe Studierender mit einem Managementspiel in unangenehme Situationen brachte – ganz unbeabsichtigt und ohne es zunächst zu bemerken. Erst als einige Studierende ihn gegen Ende der Veranstaltung darauf ansprachen, bemerkte er die negative Dynamik (ebd.). Es könnte also sein, dass Lehrende von ihren innovativen spielerischen und erfahrungsorientierten Methoden überzeugter sind als Studierende, denen sie zugemutet werden. Häufig werden Studierende das Seminar dennoch mit positiven Erlebnissen und Erkenntnissen verlassen, Elmes zeigt jedoch, dass das nicht immer der Fall ist. Für die Praxis erfordern erfahrungsorientierte Methoden daher sehr viel Sensibilität von Lehrenden und Flexibilität in der Gestaltung von Veranstaltungen. Für die Forschung eröffnet sich mit der Frage, was muten wir Lernenden mit erfahrungsorientierten Methoden zu, ein Feld, das meines Ermessens bislang wenig ausgeleuchtet ist.

Ein zweiter Kritikpunkt lässt sich sehr treffend mit folgender Aussage aus einem Experteninterview mit Allen Feldt zusammenfassen: „Games are powerful because they teach by experience. But the danger is: it teaches by experience. People are not capable of disbelieving things that happened to them.“ (Feld, Allen: zitiert nach W. C. Kriz et al., 2022). Ein Kennzeichen von Planspielen ist, dass sie ein **Modell der Realität** enthalten, mit dem Lernende sich während des Spiels auseinandersetzen (siehe Kapitel 3.1). Dieses Modell wird während des Spielens vermittelt und von Lernenden erfahrungsorientiert adaptiert – und, so die Hoffnung und zugleich Befürchtung von Feldt, geglaubt. Allerdings wird in Spielen niemals

die Realität selbst abgebildet, sondern es werden modellhafte Abstraktionen, womöglich Verzerrungen, Verkürzungen und Überbetonungen abgebildet. Es könnte ein Problem sein, wenn Spiele es tatsächlich schafften diese Zerrbilder als Realität zu vermitteln. Kriz et al. merken an, dass Planspiele per se dazu entwickelt werden, um zu beeinflussen. Sie haben pädagogische Ziele, wollen dysfunktionale Systeme darstellen und zeigen, wie sie verändert werden könnten. Jedoch sei die Grenze zwischen legitimer Beeinflussung (z.B. Bildungsabsicht) und Manipulation nicht immer klar auszumachen und teilweise nicht bewusst (2022, S. 186). Insbesondere Wirtschaftsplanspiele seien in einer Weise designt, dass sie vorwiegend neoliberale Ökonomik vermittelten, ohne dass Ethik, Ökologie und natürliche Grenzen des Wachstums im Spielmodell vorkämen (ebd., S. 189). Die Kritik von Kriz et al. lautet also, dass viele Wirtschaftsplanspiele manipulativ seien, weil sie eine bestimmte Art von neoliberalen Wirtschaftsnormen vermitteln und als gegeben darstellen. Tatsächlich reflektieren Planspielanbieter kaum, welche impliziten Annahmen über die Realität ihre Spielmodelle in sich tragen und somit vermitteln. Beispielsweise ist mir kein Handbuch zu einem Planspiel bekannt, in dem mögliche Verzerrungen und Ungenauigkeiten des Spiels explizit offengelegt werden.

Dies bringt uns nahtlos zum dritten Kritikpunkt. Ungenauigkeiten, Abstraktionen und Verzerrungen sind eigentlich eine Stärke von Planspielen und anderen erfahrungsorientierten Methoden. Ohne die gesamte Komplexität der Realität berücksichtigen zu müssen, werden wesentliche politische, soziale, ökologische oder ökonomische Systemzusammenhänge auf abstrakte Weise dargestellt und erlebbar. Relevant für die Qualität des Lernens mit erfahrungsorientierten Methoden ist jedoch die anschließende **Reflexion der Erfahrung** (Kapitel 3.1). Hierzu gehört auch die Frage inwiefern das Spielmodell mit der Realität übereinstimmt. Bereits 1999 kritisiert Reynolds (1999) im Hinblick auf die Managementausbildung, dass erfahrungsorientierte Methoden (als Beispiele nennt er Rollenspiele, Teambuilding, Simulationen und Outdoorerfahrungen) zwar häufig Anwendung finden, dass sie aber zumeist nur im Hinblick auf persönliche Kompetenzen (z.B. decision making) reflektiert werden, während übergeordnete soziale, politische, ökologische Fragen ausgespart bleiben. In Anlehnung an Giroux unterteilt Reynolds (ebd.) die Managementausbildung in Inhalte (was wird gelehrt) und Prozesse (wie wird gelehrt). Während sich innovative Prozesse (Teamwork, Simulationen, Partizipation) in der Managementausbildung

durchsetzen, würden die Inhalte auf der traditionellen Ebene verhaftet bleiben. Traditionelle Wirtschaftsinhalte (z.B. das Primat der Gewinnmaximierung) würden in spielerischen, modernen Gewändern präsentiert. Reynolds kritisiert, dass in der Managementausbildung die Idee des erfahrungsorientierten Lernens nach Kolb simplizistisch reduziert werde und entwirft ein Verständnis von kritischer Reflexion, angelehnt an Habermas und die kritische Theorie (S. 539). Für Wirtschaftsplanspiele würde Reynolds wohl vorschlagen, diese nicht vorwiegend auf der operativen Ebene auszuwerten und zu reflektieren, etwa zu fragen, wie Entscheidungen in der Simulation (und in der Realität) effektiver getroffen werden könnten, wie bestimmte Kennzahlen entstehen und wie man sie zur Steigerung des Gewinns nutzen könnte. Im Sinne der kritischen Theorie würde Reynolds stattdessen reflektieren, welche impliziten Annahmen über ein historisch konstituiertes Wirtschaftssystem die Simulation vermittelt und wie ein anderes, gerechteres Wirtschaftssystem simuliert werden könnte. Mit Reynolds könnte man Kriz et al. entgegenhalten, dass es nicht (nur) das verkürzt entworfene Planspiel ist, das ein womöglich problematisches Wirtschaftsverständnis fördert, sondern (auch) die einseitige Art der Reflexion, wobei auch Kriz et al. stark auf die Bedeutung des Debriefings hinweisen.

Dieser Exkurs soll nicht die Bedeutung von Planspielen und anderen erfahrungsorientierten Methoden schmälern, er soll vielmehr auf eher wenig beachtete Aspekte des Diskurses hinweisen.

8 Fazit und Ausblick

Diese Arbeit widmet sich der Evaluation planspielbasierter Lehrveranstaltungen. In den einleitenden Kapiteln gibt diese Arbeit daher einen Überblick über das Thema Evaluation (Kapitel 2) sowie zum Thema Planspiele in der Hochschullehre (Kapitel 3). Den Kern dieser kumulativen Arbeit bilden vier Veröffentlichungen, die in Kapitel 5 enthalten sind. Als thematische Grundlage wurde eine Literaturanalyse durchgeführt. In deren Zentrum stand die Frage welche Gelingensbedingungen zu einer erfolgreichen Lehre mit Planspielen beitragen. In den folgenden empirischen Artikeln werden drei in der Literaturanalyse identifizierte Forschungslücken aufgegriffen:

- Welche Rolle spielen Rahmenbedingungen in planspielbasierten Lehrveranstaltungen? Am Beispiel der Rahmenbedingung Online-/Präsenzlehre wurde dieses Thema für Planspiele erstmals systematisch untersucht.
- Wie kann das komplexe Lernen in Planspielen anhand unterschiedlicher Methoden sichtbar gemacht werden? Hierzu wurde in einer quasiexperimentellen Studie insbesondere untersucht, inwiefern wiederholte Reflexionsaufgaben den Lernprozess beeinflussen.
- Wie ist das Zusammenspiel unterschiedlicher Einflussfaktoren auf den Lehrerfolg und wie interagieren die unabhängigen Variablen? Mit einem Strukturgleichungsmodell wurden interdependente Zusammenhänge anhand der gesammelten Evaluationsdaten modelliert und ausgewertet.

In allen empirischen Studien kommt ein am ZMS entwickelter Fragebogen zur Evaluation planspielbasierter Lehrveranstaltungen zum Einsatz, teilweise ergänzt um weitere Forschungsmethoden. Dieser wurde ebenfalls in den einleitenden Kapiteln dieser Arbeit (Kapitel 4) vorgestellt.

Die zentralen Ergebnisse der Literaturanalyse sowie der drei empirischen Arbeiten wurden in Kapitel 6 noch einmal aufgegriffen und diskutiert. Zudem wurden Möglichkeiten der Weiterentwicklung des Fragebogens aufgezeigt, aber auch Grenzen der Evaluation komplexer Lehr-Lernmethoden beleuchtet.

Dieses finale Kapitel widmet sich der Frage, welche Schlüsse sich aus den vorhergehenden Kapiteln für die praktische Arbeit am ZMS ziehen lassen und welche Folgefragen sich für die weitere Forschung stellen.

8.1 Empfehlungen für die praktische Arbeit am ZMS

Die systematische Literaturanalyse und die Ergebnisse des Strukturgleichungsmodells weisen in dieselbe Richtung: Planspielbasierte Lehrveranstaltungen sind ein komplexes Ganzes. Sie gelingen durch das Zusammenwirken mehrerer Faktoren. Die Analyse der indirekten Effekte des Strukturgleichungsmodells macht jedoch deutlich, dass dem Lehrendenverhalten eine Schlüsselrolle zukommt. Lehrende nehmen starken Einfluss auf Faktoren, die wiederum auf den Faktor Lernen und Zufriedenheit wirken. Am ZMS sollten daher Weiterbildungen und Trainings für Lehrende weitergeführt und intensiviert werden. Das Strukturgleichungsmodell liefert auch Hinweise darauf, wie diese Trainings aussehen sollten. Nicht das Lehrendenverhalten (Facilitation) an sich nimmt direkt

starken Einfluss auf Lernen und Zufriedenheit, sondern das Planspiel mit dem Studierende interagieren. Lehrende sollten daher in der Lage sein das Planspiel und seine Möglichkeiten ins Zentrum der Lehrveranstaltung zu stellen, das Planspiel sozusagen „sprechen“ zu lassen. Dies erfordert von Lehrenden inhaltliche und methodische Kompetenzen. Sie sollten die verwendeten Planspiele sehr gut kennen und die darin enthaltenen Zusammenhänge aufzeigen können (inhaltliche Kompetenz). Zudem sollten Lehrende in der Lage sein das Seminar so zu gestalten, dass sich Studierende intensiv und reflexiv mit dem Planspiel und den darin simulierten Gegenständen befassen können (methodische Kompetenz). Neben der Expertise im Planspiel brauchen Lehrende die Methoden- und Moderationskompetenz das Planspiel tatsächlich „sprechen“ zu lassen. Im Zentrum der Trainings für Lehrende sollte die Frage stehen, wie planspielbasierte Lehrveranstaltungen so gestaltet werden können, dass Studierende sich intensiv und reflektiv mit dem Planspiel (und den darin enthaltenen Inhalten) auseinandersetzen?

The Role of Reflection in Learning with Simulation Games (Kapitel 5.3) zeigt, dass eine kurze ergänzende Reflexionsaufgabe einen Unterschied für den Lernprozess machen kann. Am ZMS sollten daher auf die jeweiligen Planspiele abgestimmte Reflexionsmethoden und -aufgaben entwickelt werden, die begleitend zum Einsatz kommen können. Hierbei ist zu beachten, dass ein wesentlicher positiver Effekt in der wiederholten Durchführung liegt. Lernende agieren im Planspiel, lassen diese Erfahrung Revue passieren, agieren wieder und erhalten noch einmal die Gelegenheit zur Reflexion usw. Debriefing sollte demnach nicht (nur) als Erklären von Spielzusammenhängen verstanden werden, sondern als eigenständiges, aktives, intellektuelles Aneignen der im Spiel aufgeworfenen Probleme und dies in mehreren Schleifen. In Kapitel 4.4 wurde ausgeführt, dass das Engagement der Studierenden (unter anderem) anhand der Frage nach der Aktivität in der Kleingruppe und anhand der Frage nach der Aktivität im Plenum (Besprechung der Spielergebnisse) operationalisiert wird. Mit einer Mittelwertdifferenz von 1.32 (sechsstufige Skala) bewerten die Studierenden ihre Aktivität im Plenum (wo vorwiegend das Debriefing stattfindet) als wesentlich geringer, als ihre Aktivität in der Kleingruppe (wo vorwiegend an der Simulation gearbeitet wird). Am ZMS sollten daher Debriefingmethoden entwickelt und in Planspiele integriert werden, die Studierende zu aktiver Reflexion anregen. Während in der Literatur aktivierende Debriefing-Methoden vorgestellt werden (Meidert, 2024), dürfte in der

Praxis die klassische Präsentation/Diskussion der Spielergebnisse vorherrschen, in der Studierende eine eher passive Rolle einnehmen. Anhand der angesprochenen Items könnte zudem überprüft werden, inwiefern Studierende ihre Aktivität bei der Besprechung der Spielergebnisse anders einschätzen, wenn das Debriefing methodischer gestaltet wird. Bei der Entwicklung von ergänzenden Debriefingmethoden könnte auch die im Exkurs (Kapitel 7) aufgeworfene Frage nach der Art der Reflexion berücksichtigt werden, sodass sich die Reflexion auch auf übergeordnete gesellschaftliche und wirtschaftliche Fragen bezieht, und nicht nur die Ebene des konkreten Managementhandelns berücksichtigt.

Ein wesentlicher Teil der empirischen Daten für dieses Projekt wurde während der Corona-Pandemie mit ihren Einschränkungen erhoben. Diese besondere Situation gab auch Anlass zum Vergleich von Präsenz- und Onlinelehrveranstaltungen (Kapitel 5.2). Die empirischen Daten zeigen, dass insbesondere computerbasierte Planspiele sehr erfolgreich als Onlineveranstaltungen umgesetzt werden können. Die Zeit der pandemiebedingten Onlinelehre ist (glücklicherweise) vorbei. Einflussfaktoren wie Streiks oder hochschulübergreifende (internationale) Projekte könnten es dennoch erforderlich oder sinnvoll machen, Onlineveranstaltungen anzubieten. Die Evaluationsdaten weisen darauf hin, dass die didaktischen Konzepte hierfür am ZMS vorhanden sind und ohne nennenswerte Qualitätseinschränkung durchgeführt werden können.

Am ZMS wurden erstmals planspielbasierte Lehrveranstaltungen in großer Zahl evaluiert. Ziel zu Beginn des Projekts war es, möglichst viele Planspiele in möglichst unterschiedlichen Settings (z.B. unterschiedliche Lehrende, Onlinelehre, Präsenzlehre, Kursgröße) zu evaluieren. Mit über 3000 Fällen aus 172 Veranstaltungen ist nun ein guter Überblick vorhanden, der 31 unterschiedliche Planspiele und unterschiedliche Durchführungsformen einschließt. Dennoch sind für eher selten eingesetzte Planspiele nur wenige Fälle vorhanden, während für häufig eingesetzte Planspiele sehr viele Veranstaltungen evaluiert wurden. Für Planspiele, die bereits in hoher Zahl evaluiert wurden, ist nicht zu erwarten, dass sich (bei ähnlicher Art der Durchführung) neue Erkenntnisse ergeben werden. Die Evaluationsstrategie könnte daher in zweifacher Hinsicht angepasst werden. Während Veranstaltungen mit häufig evaluierten Planspielen nur noch stichpunktartig evaluiert werden, könnten insbesondere wenig eingesetzte Planspiele fokussiert werden. Zudem könnten in der Evaluation thematische Schwerpunkte gesetzt werden. Wie in Kapitel 6.5 diskutiert, ist es in der

Standardevaluation kaum möglich alle Aspekte des komplexen Lehr-Lernprozesses abzudecken. Jedoch könnte die Standardevaluation für einige Veranstaltungen im Hinblick auf einen thematischen Schwerpunkt (z.B. Erlernen von Fachkompetenzen) ergänzt werden. Während zu Beginn der strukturierten Evaluation am ZMS eine möglichst große Breite an Veranstaltungen evaluiert wurde, um einen guten Überblick zu erhalten, könnte die Strategie so angepasst werden, dass künftig eher spezifischere Fragen im Vordergrund stehen (z.B. bislang wenig evaluierte Planspiele/Durchführungsformen oder die Evaluation im Hinblick auf bestimmte Fragestellungen). Schwerpunkt der Evaluationsstrategie wäre also nicht mehr eine möglichst große Breite an Veranstaltungen zu erheben, sondern spezifische Charakteristika in den Fokus zu nehmen.

8.2 Ansätze für die weitere Forschung

Das Fazit des Literatur Reviews wirft ein bislang unbearbeitetes theoretisch-konzeptionelles Thema auf: „Die große Breite der Forschung bringt auch Schwierigkeiten mit sich. Es scheint international unklar zu sein, was ein Planspiel eigentlich ist. Während im deutschsprachigen Raum häufig die Definition von Kriz (2011) aufgegriffen wird, scheint sich im englischsprachigen Raum keine Definition durchgesetzt zu haben. So werden unter dem Begriff „simulation game“ Computerspiele, Lernspiele und Planspiele gleichermaßen diskutiert“ (Alf, 2022, S. 475). Insbesondere die internationale englischsprachige Literatur könnte auf den Gehalt des Begriffs „simulation game“ hin untersucht werden, mit dem Ziel spezifischere Definitionen für unterschiedliche Arten von Spielen und Durchführungsformen zu entwickeln.

In mehreren Teilen dieser Arbeit wurde die Bedeutung der Lehrenden für die erfolgreiche Durchführung von Planspielen hervorgehoben. Die Literaturstudie fasst zusammen: „Die Bedeutung der Lehrenden in Planspielen ist zentral. Sie moderieren zwischen Simulation und Lernenden. Erst durch Einführung, Anpassung, Erklärung, Feedback und schlussendlich die Auswertung der Simulation entsteht für Lernende ein Mehrwert“ (Alf, 2022, S. 475). Die Arbeit mit den empirischen Daten bestätigen die zusammenfassende Erkenntnis der ausgewerteten Literatur: Facilitation nimmt den größten Einfluss auf Lernen und Zufriedenheit, insbesondere vermittelt durch Verständnis und Relevanz des Planspiels (Kapitel 5.4). Dieses Thema sollte in der Forschung weiter beleuchtet werden. Welches Lehrverhalten und welche didaktisch-/methodische

Vorgehensweisen werden als hilfreich für das Verständnis und schlussendlich für das Lernen angesehen?

Nach wie vor wird die Abfolge von Briefing, Spielerfahrung und Debriefing als typischer Ablauf planspielbasierter Lehrveranstaltungen beschrieben (Rappenglück, 2017). Schwägele et al. schlagen jedoch, vor diesen Ablauf aufzuweichen und das Verhältnis von Erfahrung und Debriefing bzw. Reflexion der Erfahrung fluide zu gestalten (Schwägele et al., 2021). Der Beitrag *The Role of Reflection in Learning with Simulation Games* untersucht auch, inwiefern wiederkehrende Reflexionsaufgaben, also ein wiederholtes Unterbrechen der Spielerfahrung, förderlich sind. Besonders die qualitative Auswertung im Rahmen von Forschungsfrage 4 (Alf et al., 2023, S. 628) macht das Potential wiederholter Reflexion deutlich: Während in früheren Runden relevante Probleme identifiziert werden, können sie in späteren Runden erfolgreich bearbeitet werden. In Zukunft sollte also von der starren Aufteilung in Briefing, Erfahrung und Debriefing abgesehen werden und eher wiederholende, fluide Modelle der Planspielreflexion entwickelt werden. Zudem untersucht *The Role of Reflection* den Unterschied zwischen einer klassischen Debriefingmethode (gemeinsame Besprechung der Spielergebnisse, Erklärung von Zusammenhängen und die Möglichkeit, Fragen zu stellen) mit dem klassischen Debriefing ergänzt um eine reflektive Gruppenaufgabe. Die Ergebnisse weisen darauf hin, dass die Reflexionsaufgabe zu wesentlichen Änderungen führt. Daraus kann abgeleitet werden, dass das Debriefing methodisch so gestaltet werden sollte, dass Lernende sich selbst reflexiv mit den gesammelten Erfahrungen auseinandersetzen. Obwohl die Rolle der Reflexion in Lernprozessen mit Planspielen häufig betont wird, fällt die empirische Erforschung dieser Prozesse eher dünn aus. Zukünftige Forschung sollte weiter untersuchen, inwiefern Reflexionsprozesse beim Lernen eine Rolle spielen und mit welchen didaktischen Methoden Reflexionsprozesse ermöglicht werden können.

Auch die Rolle der Teamarbeit in Planspielen sollte weiter untersucht werden. Die ausgewertete Literatur (Kapitel 5.1) weist auf Zusammenhänge zwischen guter Teamarbeit und Lernen mit Planspielen hin. In der Pfadanalyse hingegen nimmt die Variable Funktion innerhalb des Teams keinen Einfluss auf Lernen und Zufriedenheit (Kapitel 5.4, RQ1). Die in der Literaturanalyse ausgewerteten Studien erheben das Lernen mit Planspielen vorwiegend anhand des Erfolgs im Planspiel (Tao et al., 2012). Evtl. nimmt die Art der Teamarbeit stärkeren Einfluss

auf die Performance, jedoch nur in geringer Weise auf das selbst eingeschätzte Lernen. Der Einfluss von Teamarbeit auf Lernen einerseits und Performance andererseits könnte anhand von Spieldaten, Lerntests und Selbsteinschätzungen zum Lernen in einer eher experimentell angelegten Studie weiter untersucht werden.

In der Diskussion zur Weiterentwicklung des Fragebogens (Kapitel 6.4) wurde angedeutet, dass es einen Unterschied machen könnte, ob ein Planspiel als Teil einer größeren Lehrveranstaltung stattfindet, in der weitere Themen und Methoden Anwendung finden – oder ob es als abgeschlossene Lehreinheit in Form einer Blockveranstaltung an einem spezialisierten Zentrum stattfindet. Diese Unterscheidung ist keine Besonderheit von Planspielen. Auch weitere Themen werden an Hochschulen sowohl als Teil „normaler“ Lehrveranstaltungen behandelt und andererseits an spezifischen Zentren gelehrt. So gibt es beispielsweise an vielen Universitäten und Hochschulen Zentren, die spezielle Veranstaltungen für wissenschaftliches Schreiben anbieten¹⁸, während wissenschaftliches Schreiben auch Teil einer „normalen“ Veranstaltung zum wissenschaftlichen Arbeiten sein kann. Anhand dieser Beispiele stellt sich die Frage, welche Vor- und Nachteile sich für Studium und Lehre ergeben, wenn für bestimmte Themen oder Methoden spezialisierte Zentren gebildet werden. Diese Fragestellung könnte anhand verschiedener Beispiele (z.B. Planspiele, wissenschaftliches Schreiben, empirische Methoden) untersucht werden.

¹⁸ Als Beispiele seien hier das Schreibzentrum der Goethe Universität Frankfurt (https://www.starkerstart.uni-frankfurt.de/82720027/PortalStartPage_82720027) und das Zentrum für Wissenschaftliches Schreiben an der FH Wien (<https://www.fh-campuswien.ac.at/lehre/zentrum-fuer-wissenschaftliches-schreiben.html>) genannt.

Danke!

[Danksagungen wurden aus der digitalen Version entfernt]

Literatur (Rahmentext)

- Al-Elq, A. H. (2010). Simulation-based medical teaching and learning. *Journal of family & community medicine*, 17(1), 35–40. <https://doi.org/10.4103/1319-1683.68787>
- Alf, T. (2022). Gelingensbedingungen von Planspiellehrveranstaltungen – Ein Systematic Literature Review. *die hochschullehre*, 8(33), 468–480. <https://doi.org/10.3278/HSL2233W>
- Alf, T. & Trautwein, F. (2023). Planspielbasierte Lehrveranstaltungen: Ein Vergleich zwischen Präsenz- und Onlinelehre. *Zeitschrift für Hochschulentwicklung*, 18, 345–363.
- Alf, T., Wijse, M. de & Trautwein, F. (2023). The Role of Reflection in Learning with Simulation Games – A Multi-Method Quasi Experimental Research. *Simulation & Gaming*, 54(6), 621–644. <https://doi.org/10.1177/10468781231194896>
- Arbeitskreis „Lehrevaluation“ im Fach Psychologie, Zentrum für Psychologische Diagnostik, Begutachtung und Evaluation, Gläßer, E., Gollwitzer, M [M.], Kranz, D., Meiniger, C., Schlotz, W., Schnell, T. & Voß, A. (2002). *TRIL - Trierer Inventar zur Lehrevaluation*. <https://doi.org/10.23668/psycharchives.6590>
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2018). *Multivariate Analysemethoden*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-56655-8>
- Bartschat, D., Zürn, B., Trautwein, F. & Hühn, C. (2018), Planspielseminare zum Studieneinstieg – die Erstsemesterformate im Zentrum für Managementsimulation an der DHBW Stuttgart. In D. Ternes & C. C. Schnekenburger (Hrsg.), *#DUAL: Band 1. Facetten der Lehre* (S. 101–110). Duale Hochschule Baden-Württemberg.
- Beranic, T. & Hericko, M. (2019). Introducing ERP Concepts to IT Students Using an Experiential Learning Approach with an Emphasis on Reflection. *Sustainability*, 11(18), 1–17. <https://doi.org/10.3390/su11184992>
- Berning, C. C. (2019). Strukturgleichungsmodelle. In M. Apelt, I. Bode, R. Hasse, U. Meyer, V. V. Groddeck, M. Wilkesmann & A. Windeler (Hrsg.), *Springer reference Sozialwissenschaften. Handbuch Organisationssoziologie* (S. 1–18). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-16937-4_30-2
- Bitterer, F., Frank, A. & Weiß, P. Lernzielorientierte Evaluation von Lehrveranstaltungen – das Bielefelder Modell (BiLOE). In M. Fuhrmann, J. Güdler, P. Pohlenz & U. Schmidt (Hrsg.), *Handbuch Qualität in Studium, Lehre und Forschung*. [https://www.hqsl-bibliothek.de/de/handbuch/gliederung/#/Beitragsdetailansicht/449/2833/Lernzielorientierte-Evaluation-von-Lehrveranstaltungen-%25E2%2580%2593-das-Bielefelder-Modell-\(BiLOE\)](https://www.hqsl-bibliothek.de/de/handbuch/gliederung/#/Beitragsdetailansicht/449/2833/Lernzielorientierte-Evaluation-von-Lehrveranstaltungen-%25E2%2580%2593-das-Bielefelder-Modell-(BiLOE))
- Blötz, U. (2003). *Planspiele in der beruflichen Bildung: Aktualisierter Planspielkatalog und neue Fachbeiträge 2003 : Abriss zur Auswahl, Konzeptionierung und Anwendung von Planspielen : Multimedia-Publikation* (3. überarb. Aufl.). Bertelsmann.

- Blötz, U. (Hrsg.). (2015). *Berichte zur beruflichen Bildung. Planspiele und Serious Games in der beruflichen Bildung: Auswahl, Konzepte, Lernarrangements, Erfahrungen - aktueller Katalog für Planspiele und Serious Games 2015* (5., überarbeitete Auflage). BIBB Bundesinstitut für Berufsbildung; W. Bertelsmann Verlag GmbH & Co. KG.
- Bortz, J. & Döring, N. (Hrsg.). (2006). *Springer-Lehrbuch. Forschungsmethoden und Evaluation*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-33306-7>
- Bortz, J. & Schuster, C. (2010). *Statistik: Für Human- und Sozialwissenschaftler : mit 163 Tabellen* (7., vollst. überarb. und erw. Aufl.). Springer.
- Böttcher, W., Holtappels, H. G. & Brohm, M. (2006). Evaluation im Bildungswesen. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Grundlagentexte Pädagogik. Evaluation im Bildungswesen: Eine Einführung in Grundlagen und Praxisbeispiele* (S. 7–22). Juventa-Verl.
- Braun, E., Gusy, B., Leidner, B. & Hannover, B. (2008). Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). *Diagnostica*, 54(1), 30–42. <https://doi.org/10.1026/0012-1924.54.1.30>
- Brennan, R. & Vos, L. (2013). Effects of Participation in a Simulation Game on Marketing Students' Numeracy and Financial Skills. *Journal of Marketing Education*, 35(3), 259–270. <https://doi.org/10.1177/0273475313482928>
- Campos, N., Nogal, M., Caliz, C. & Juan, A. A. (2020). Simulation-based education involving online and on-campus models in different European universities. *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-0181-y>
- Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn? *Journal of Marketing Education*, 31(1), 16–30. <https://doi.org/10.1177/0273475308324086>
- Costin, Y., O'Brien, M. P. & Slattery, D. M. (2018). Using Simulation to Develop Entrepreneurial Skills and Mind-Set: An Exploratory Case Study. *International Journal of Teaching and Learning in Higher Education*, 30(1), 136–145. <https://www.isetl.org/ijtlhe/pdf/IJTLHE2802.pdf>
- Dash, G. & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, 173, 1–11. <https://doi.org/10.1016/j.techfore.2021.121092>
- Decristan, J., Hess, M., Holzberger, D. & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale: Eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. In A.-K. Praetorius & E. Klieme (Hrsg.), *Zeitschrift für Pädagogik. Beiheft: Bd. 66. Empirische Forschung zu Unterrichtsqualität: Theoretische Grundfragen und quantitative Modellierungen* (1. Auflage, S. 102–116). Beltz Juventa.
- Deterding, S., Dixon, D., Khaled, R. & Nacke, L. From game design elements to gamefulness. In *Lugmayr, Franssila et al. (Hg.) – Proceedings of the 15th International* (S. 9–15). <https://doi.org/10.1145/2181037.2181040>
- Deutsche Gesellschaft für Evaluation e.V. (o.D.). *Glossar der Standards für Evaluation*. Deutsche Gesellschaft für Evaluation e.V. <https://www.degeval.org/degeval-standards/glossar-der-standards-fuer-evaluation/>

- Dörner, R., Göbel, Stefan, Effelsberg, Wolfgang & Wiemeiyer, J. (2016). *Serious Games: Foundations, Concepts and Practice* (1st ed. 2016). Springer International Publishing.
- Duke, R. D. & Geurts, J. (2004). *Policy games for strategic management: Pathways to the Unknown*. Dutch University Press.
- Ehrenspeck, Y., Haan, G. de & Thiel, F. (Hrsg.). (2008). *Bildung: Angebot oder Zumutung?* VS Verlag für Sozialwissenschaften.
- Eid, M., Gollwitzer, M [Mario] & Schmitt, M. (2013). *Statistik und Forschungsmethoden: Lehrbuch ; mit Online-Materialien* (3., korr. Aufl.). Beltz.
- Elmes, M. B. (2019). Working With(in) the Shadow of Experiential Learning. *Journal of Management Education*, 43(1), 99–107.
<https://doi.org/10.1177/1052562918804587>
- Engartner, T. & Meßner, M. T. (2017). Tarifverhandlungen als Gegenstand von Planspielen in der arbeitnehmer- und gewerkschaftsorientierten Bildung. In A. Petrik & S. Rappenglück (Hrsg.), *Reihe Politik und Bildung: Band 81. Handbuch Planspiele in der politischen Bildung* (S. 69–76). Wochenschau Verlag.
- Entner, C., Fleischmann, A. & Strasser, A. (2021). Hochschullehre im digitalen Wandel. Überlegungen zur didaktischen Gestaltung von Präsenz- und Onlinelehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba, M. Wiemer & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre*. DUZ Verlags- und Medienhaus GmbH.
- EvaSys. (o.J.). *PDF-Report Plugin Lernerfolgsevaluation*. https://webar-chiv.typo3.tum.de/WI/leist/fileadmin/w00but/www/PDF-Report_Plugin_Lernerfolgsevaluation.pdf
- Freese, M. (2019). *Mensch frustrierte dich nicht oder doch? Der Einfluss von Emotionen auf Ergebnisfaktoren von Team-Entscheidungen am Beispiel eines Planspieles für das Flughafenmanagement* [Dissertation]. Karlsruher Institut für Technologie (KIT), Karlsruhe.
- Gäde, J. C., Schermelleh-Engel, K. & Brandt, H. (2020). Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 615–659). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-662-61532-4_24
- Geithner, S. & Menzel, D. (2016). Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. *Simulation & Gaming*, 47(2), 228–256. <https://doi.org/10.1177/1046878115624312>
- Geuting, M. (1992). *Planspiel und soziale Simulationen im Bildungsbereich. Studien zur Pädagogik, Andragogik und Gerontagogik: Bd. 10*. Verlag Peter Lang.
- Geuting, M. (2000). Soziale Simulation und Planspiel in pädagogischer Perspektive. In D. Herz & A. Blättle (Hrsg.), *Grundlegung und Methoden der Politischen Wissenschaft. Simulation und Planspiel in den Sozialwissenschaften: Eine Bestandsaufnahme der internationalen Diskussion* (S. 15–62). Lit Verlag.
- Gnahn, D. & Quilling, E. (Hrsg.). (2019). *Qualitätsmanagement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-19534-2>

- Hair, J. F., Risher, J. J., Sarstedt, M. & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hayes, A. F. & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Hense, J., Böttcher, W., Kalman, M. & Meyer, W. (Hrsg.). (2019). *Evaluation: Standards in unterschiedlichen Handlungsfeldern: Einheitliche Qualitätsansprüche trotz heterogener Praxis?* [1. Auflage]. Waxmann.
- Hense, J., Kriz, W. C. & Wolfe, J. (2009). Putting Theory-Oriented Evaluation Into Practice. *Simulation & Gaming*, 40(1), 110–133. <https://doi.org/10.1177/1046878107308078>
- Hettler, I. S., Rahn, S. & Meyer, T. (2022). *Dual studieren unter erschwerten Bedingungen – wie bewältigen Studierende in belastenden Lebenslagen ihr Studium an der DHBW während der COVID-19-Pandemie. Eine explorative Interview-Studie.* DHBW. [https://www.dhbw.de/fileadmin/user_upload/Dokumente/Schrifterzeugnisse/Dual studieren unter erschwerten Bedingungen.pdf](https://www.dhbw.de/fileadmin/user_upload/Dokumente/Schrifterzeugnisse/Dual_studieren_unter_erschwertten_Bedingungen.pdf)
- Ijsselsteijn, W. A., Kort, Y. de & Poels, K. (2013). *The Game Experience Questionnaire.* Eindhoven University of Technology. <https://research.tue.nl/en/publications/the-game-experience-questionnaire>
- The jamovi project. (2023). *jamovi* (Version 2.3) [Computer software]. <https://www.jamovi.org>
- Kadel, J., Buschmann, C., Haas, S., Meßner, M. T. & Adl-Amini, K. (2023). Planspiele und simulative Methoden in der Lehrkräftebildung – ein Literaturüberblick. *Zeitschrift für Hochschulentwicklung*, 18, 19–39.
- Kawalle, S. (2017). Planspiele: Lehren und Forschen in Synergie. In A. Petrik & S. Rappenglück (Hrsg.), *Reihe Politik und Bildung: Band 81. Handbuch Planspiele in der politischen Bildung* (S. 181–202). Wochenschau Verlag.
- Klabbers, J. H. G. (2018). On the Architecture of Game Science. *Simulation & Gaming*, 49(3), 207–245. <https://doi.org/10.1177/1046878118762534>
- Knödler, E. (2019). *Evaluation an Hochschulen: Entwicklung und Validierung eines verhaltensbasierten Messinventars zur studentischen Lehrveranstaltungsevaluation.* Springer VS. <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5716836>
- Koopmann, J., Zimmer, L. M. & Lörz, M. (2024). The impact of COVID-19 on social inequalities in German higher education. An analysis of dropout intentions of vulnerable student groups. *European Journal of Higher Education*, 14(2), 290–307. <https://doi.org/10.1080/21568235.2023.2177694>
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomie: An Overview. *Theory into Practice*, 41(4). <https://www.depauw.edu/files/resources/krathwohl.pdf>
- Kriz, W. (2009). Planspiel. In S. Kühl, P. Strodtholz & A. Taffertshofer (Hrsg.), *Handbuch Methoden der Organisationsforschung* (S. 558–578). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91570-8_27
- Kriz, W. & Manahl, W. (2016). Design of Simulation Games for Systems Management Education. In S. Schwägele, B. Zürn, D. Bartschat & F. Trautwein

- (Hrsg.), *ZMS-Schriftenreihe: Band 8. Planspiele - Vernetzung gestalten: Forschungsergebnisse und Praxisbeispiele für morgen* (S. 69–92). Books on Demand.
- Kriz, W. C. (2010). A Systemic-Constructivist Approach to the Facilitation and Debriefing of Simulations and Games. *Simulation & Gaming*, 41(5), 663–680. <https://doi.org/10.1177/1046878108319867>
- Kriz, W. C., Kikkawa, T. & Sugiura, J. (2022). Manipulation Through Gamification and Gaming. In T. Kikkawa, W. C. Kriz & J. Sugiura (Hrsg.), *Translational Systems Sciences. Gaming as a Cultural Commons* (Bd. 28, S. 185–199). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-0348-9_11#DOI
- Kriz, W. C. & Manahl, W. (2018). Gaming Simulation as a Science of Design Approach. In A. Naweed, M. Wardaszko, E. Leigh & S. Meijer (Hrsg.), *Lecture Notes in Computer Science. Intersections in Simulation and Gaming* (Bd. 10711, S. 380–393). Springer International Publishing. https://doi.org/10.1007/978-3-319-78795-4_27
- Kromrey, H. (2005). Evaluation - ein Überblick. In H. Schöch (Hrsg.), *Schriftenreihe Wandel und Kontinuität in Organisationen: Bd. 6. Was ist Qualität? Die Entzauberung eines Mythos* (S. 31–85).
- Lehner, M. & Sohm, K. (2021). Qualität, didaktische Methodik und Digitalität: Erfahrungen und Reflexionen an der Fachhochschule Technikum Wien. In U. Dittler & C. Kreidl (Hrsg.), *Wie Corona die Hochschullehre verändert* (S. 339–350). Springer Fachmedien Wiesbaden.
- Magylaité, K., Kapočius, K., Butleris, R. & Čeponienė, L. (2022). Towards High Usability in Gamified Systems: A Systematic Review of Key Concepts and Approaches. *Applied Sciences*, 12(16), 8188. <https://doi.org/10.3390/app12168188>
- Mandl, H. (2004). On the path to a new learning culture - the contribution of web-based business games. In T. Eberle (Hrsg.), *Bridging the gap: transforming knowledge into action through gaming and simulation: Proceedings of the 35th Conference of the International Simulation and Gaming Association, Munich, 2004* (S. 163–174).
- MARSH, H. W. (1982). SEEQ: A RELIABLE, VALID, AND USEFUL INSTRUMENT FOR COLLECTING STUDENTS' EVALUATIONS OF UNIVERSITY TEACHING. *British Journal of Educational Psychology*, 52(1), 77–95. <https://doi.org/10.1111/j.2044-8279.1982.tb02505.x>
- Matute-Vallejo, J. & Melero-Polo, I. (2019). Understanding Online Business Simulation Games: The Role of Flow Experience, Perceived Enjoyment and Personal Innovativeness. *Australasian Journal of Educational Technology*, 35(3), 71–85. <https://doi.org/10.14742/ajet.3862>
- Meadows, D. H. (2009). *Thinking in systems: A primer*. Earthscan.
- Meidert, N. (2024). Spielerisches Debriefing. In T. Alf, S. Hahn, I. Fischer, B. Zürn & F. Trautwein (Hrsg.), *ZMS-Schriftenreihe: Bd. 14. Planspiele – interdisziplinär vernetzt. Rückblick auf das 34. Europäische Planspielforum und den Deutschen Planspielpreis 2023* (S. 20–28). ZMS.

- Monat, J. P. & Gannon, T. F. (2015). What is Systems Thinking? A Review of Selected Literature Plus Recommendations. *American Journal of Systems Science*, 4(1), 11–26. https://mycourses.aalto.fi/pluginfile.php/1187882/mod_label/intro/Monat%20%20Gannon_2015_What%20is%20systems%20thinking%20-%20A%20review%20of%20selected%20literature%20plus%20recommendations.pdf
- Müller, C. E. (2017). Kausale Wirkungsevaluation zwischen methodischem Anspruch und empirischer Praxis. In R. Stockmann & W. Meyer (Hrsg.), *Sozialwissenschaftliche Evaluationsforschung: Band 13. Die Zukunft der Evaluation: Trends, Herausforderungen, Perspektiven* (1. Auflage, S. 205–222) [Verlag nicht ermittelbar].
- Ney, M., Gonçalves, C. & Balacheff, N. (2014). Design Heuristics for Authentic Simulation-Based Learning Games. *IEEE Transactions on Learning Technologies*, 7(2), 132–141. <https://doi.org/10.1109/TLT.2014.2316161>
- Peters, V. A. M. & Vissers, G. A. N. (2004). A Simple Classification Model for Debriefing Simulation Games. *Simulation & Gaming*, 35(1), 70–84. <https://doi.org/10.1177/1046878103253719>
- Petranek, C. F., Corey, S. & Black, R. (1992). Three Levels of Learning in Simulations: Participating, Debriefing, and Journal Writing. *Simulation & Gaming*, 23(2), 174–185. <https://doi.org/10.1177/1046878192232005>
- Petri, G., Gresse von Wangenheim, C. & Borgatto, A. F. (2017). MEEGA+, Systematic Model to Evaluate Educational Games. In N. Lee (Hrsg.), *Encyclopedia of Computer Graphics and Games* (S. 1–7). Springer International Publishing. https://doi.org/10.1007/978-3-319-08234-9_214-1
- Petrik, A. & Rappenglück, S. (Hrsg.). (2017). *Reihe Politik und Bildung: Band 81. Handbuch Planspiele in der politischen Bildung*. Wochenschau Verlag.
- Plass, J. L., Homer, B. D., Kinzer, C. K., Chang, Y. K., Frye, J., Kaczetow, W., Isbister, K. & Perlin, K. (2013). Metrics in Simulations and Games for Learning. In M. Seif El-Nasr, A. Drachen & A. Canossa (Hrsg.), *Game Analytics* (S. 697–729). Springer London. https://doi.org/10.1007/978-1-4471-4769-5_31
- Pohlenz, P. (2008a). *Datenqualität als Schlüsselfrage der Qualitätssicherung von Lehre und Studium an Hochschulen. Potsdamer Beiträge zur Lehrevaluation: Bd. 3*. Universitätsverl.
- Pohlenz, P. (2008b). Lehrevaluation und Qualitätsmanagement: Neue Anforderungen für die Hochschulsteuerung. *Sozialwissenschaften und Berufspraxis*, 31(1), 66–78.
- Rappenglück, S. (2017). Planspiele in der Praxis der politischen Bildung: Entwicklung, Durchführung, Varianten und Trends. In A. Petrik & S. Rappenglück (Hrsg.), *Reihe Politik und Bildung: Band 81. Handbuch Planspiele in der politischen Bildung* (S. 17–34). Wochenschau Verlag.
- Reinmann, G. (2016). Gestaltung akademischer Lehre: semantische Klärungen und theoretische Impulse zwischen Problem- und Forschungsorientierung. *Zeitschrift für Hochschulentwicklung*, 11(5), 225–244.

- Reynolds, M. (1999). Critical Reflection and Management Education: Rehabilitating Less Hierarchical Approaches. *Journal of Management Education*, 23(5), 537–553. <https://doi.org/10.1177/105256299902300506>
- Riegger, M. J. (2023). Planspiele an der Hochschule in der Theologie – ein problemgeschichtlich-systematischer Überblick. *Zeitschrift für Hochschulentwicklung*, 18(Sonderheft Planspiele), 61–78. <https://doi.org/10.21240/zfhe/SH-PS/04>
- Rindermann, H. (2003). Methodik und Anwendung der Lehrveranstaltungsevaluation für die Qualitätsentwicklung an Hochschulen. *Sozialwissenschaften und Berufspraxis (SuB)*, 26(4).
- Rindermann, H. (2009). *Lehrevaluation: Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen mit einem Beitrag zur Evaluation computerbasierter Unterrichts* [2., leicht korrigierte Aufl.]. *Psychologie: Bd. 42*. Verl. Empirische Pädagogik.
- Schmidt, U. & Pohlenz, P. (2019). Evaluationsstandards im Handlungsfeld Hochschule. In J. Hense, W. Böttcher, M. Kalman & W. Meyer (Hrsg.), *Evaluation: Standards in unterschiedlichen Handlungsfeldern: Einheitliche Qualitätsansprüche trotz heterogener Praxis?* [1. Auflage], S. 137–150. Waxmann.
- Schwägele, S. (2013). *Planspiel versus Simulation? Versuch einer Abgrenzung*. Zentrum für Managementsimulation. <https://zms.dhbw-stuttgart.de/artikel/planspiel-versus-simulation-versuch-einer-abgrenzung/>
- Schwägele, S. (2015). *Planspiel - Lernen - Lerntransfer: Eine subjektorientierte Analyse von Einflussfaktoren*. Dissertation. ZMS-Schriftenreihe: Band 7 [348 Seiten].
- Schwägele, S., Zürn, B., Lukosch, H. K. & Freese, M. (2021). Design of an Impulse-Debriefing-Spiral for Simulation Game Facilitation. *Simulation & Gaming*, 52(3), 364–365. <https://doi.org/10.1177/10468781211006752>
- Spinath & Seifried (2018). Was brauchen wir, um solide empirische Erkenntnisse über gute Hochschullehre zu erhalten? *Zeitschrift für Hochschulentwicklung* (Jg.13/Nr.1), Artikel 8, 153–169. <https://doi.org/10.3217/zfhe-13-01/08>
- Spooren, P., Brockx, B. & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching. *Review of Educational Research*, 83(4), 598–642. <https://doi.org/10.3102/0034654313496870>
- Stockmann, R. (2006). Qualitätsmanagement und Evaluation im Vergleich. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Grundlagentexte Pädagogik. Evaluation im Bildungswesen: Eine Einführung in Grundlagen und Praxisbeispiele* (S. 23–38). Juventa-Verl.
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tao, Y.-H., Yeh, C. R. & Hung, K. C. (2012). Effects of the heterogeneity of game complexity and user population in learning performance of business simulation games. *Computers & Education*, 59(4), 1350–1360. <https://doi.org/10.1016/j.compedu.2012.06.003>
- Taylor, A.-S. A., Backlund, P. & Niklasson, L. (2012). The Coaching Cycle. *Simulation & Gaming*, 43(5), 648–672. <https://doi.org/10.1177/1046878112439442>

- Tosey, P., Dhaliwal, S. & Hassinen, J. (2015). The Finnish Team Academy model: Implications for management education. *Management Learning*, 46(2), 175–194. <https://doi.org/10.1177/1350507613498334>
- Trautwein, C. (2011). *Unternehmensplanspiele im industriebetrieblichen Hochschulstudium: Analyse von Kompetenzerwerb, Motivation und Zufriedenheit am Beispiel des Unternehmensplanspiels TOPSIM - General Management II*. Zugl.: Hohenheim, Univ., Diss., 2010 (1. Aufl.). Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden. <http://dx.doi.org/10.1007/978-3-8349-6955-2> <https://doi.org/10.1007/978-3-8349-6955-2>
- Trautwein, F. & Alf, T. (2023). Theory-Based Development of an Inventory for the Evaluation of Simulation Game Lectures. In C. Harteveld, S. Sutherland, G. Troiano, H. Lukosch & S. Meijer (Hrsg.), *Simulation and Gaming for Social Impact* (Bd. 13622, S. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-031-37171-4_1
- Trautwein, F. & Alf, T. (2024). Die Rolle der Lehrenden im Kontext von Planspielveranstaltungen. In T. Alf, S. Hahn, I. Fischer, B. Zürn & F. Trautwein (Hrsg.), *ZMS-Schriftenreihe: Bd. 14. Planspiele – interdisziplinär vernetzt. Rückblick auf das 34. Europäische Planspielforum und den Deutschen Planspielpreis 2023* (S. 44–57). ZMS.
- Treen, E., Atanasova, C., Pitt, L. & Johnson, M. (2016). Evidence From a Large Sample on the Effects of Group Size and Decision-Making Time on Performance in a Marketing Simulation Game. *Journal of Marketing Education*, 38(2), 130–137. <https://doi.org/10.1177/0273475316653433>
- Weiber, R. & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-35012-2>
- Wijse-van Heeswijk, M. de. (2023). Facilitation Interventions to Increase Learning Effectiveness in Game Simulations. A Generic Approach of Facilitation Applicable to a Broad Variety of Simulation Games. In M. L. Angelini & R. Muñiz (Hrsg.), *Springer Texts in Education. Simulation for Participatory Education* (S. 53–85). Springer International Publishing. https://doi.org/10.1007/978-3-031-21011-2_4
- Xu, D. & Xu, Y. (2020). The Ambivalence About Distance Learning in Higher Education. In L. W. Perna (Hrsg.), *Higher Education: Handbook of Theory and Research. Higher Education: Handbook of Theory and Research* (Bd. 35, S. 351–401). Springer International Publishing. https://doi.org/10.1007/978-3-030-31365-4_10
- Zeiner-Fink, S [Susan], Bullinger, A. C. & Geithner, S. (2023). Learning Effects and Acceptance in Business Games: A Systematic Literature Review. In C. Harteveld, S. Sutherland, G. Troiano, H. Lukosch & S. Meijer (Hrsg.), *Simulation and Gaming for Social Impact* (S. 36–51). Springer International Publishing.
- Zeiner-Fink, S [Susann], Bullinger, A. C. & Geithner, S. (2023). Learning Effects and Acceptance in Business Games: A Systematic Literature Review. In C. Harteveld, S. Sutherland, G. Troiano, H. Lukosch & S. Meijer (Hrsg.), *Simulation and Gaming for Social Impact* (Bd. 13622, S. 36–51). Springer International Publishing. https://doi.org/10.1007/978-3-031-37171-4_3

Zürn, B., Trautwein, F. & Freese, M. (2023). Empfehlungen zur curricularen Integration von Planspielen in Wirtschafts-Studiengänge. *Zeitschrift für Hochschulentwicklung*, 18, 176–185. <https://www.zfhe.at/index.php/zfhe/article/view/1725>

9 Anhang

Angenommene Faktorenstruktur Pretest

Kategorie	Quelle	Item
Einstieg	Knödler	Mein Interesse am Planspiel war zu Beginn der Veranstaltung...
Gesamtbewertung	TRIL	Alles in allem hat sich der Besuch dieser Veranstaltung für mich gelohnt.
	Trautwein	Die Teilnahme am Planspiel hat sich für mich gelohnt.
		Die Teilnahme am Planspiel war für mich zufriedenstellend.
		Ich habe beim Planspiel viel gelernt.
		Das Planspiel war für mich sehr motivierend.
	Knödler	Ich würde die Teilnahme an dieser Lehrveranstaltung anderen Studierenden weiterempfehlen.
		Mein Interesse am Planspiel war zum Ende der Veranstaltung...
		Welche Schulnote würden Sie der Veranstaltung insgesamt geben?
	DHBW LVE	Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?
	Lernen/Kompetenzgewinn	Trautwein
Ich habe beim Planspiel viel gelernt.		
Ich habe mehr gelernt als in einer normalen Vorlesung.		
Die Intensität des Lernens war hoch.		
TRIL		Das Thema der Veranstaltung hat mich interessiert.
		Die behandelten Themen waren für mich bedeutsam und relevant.
		Ich habe in dieser Veranstaltung etwas Sinnvolles und Wichtiges gelernt.
		Mein Verständnis für das Studienfach hat sich durch die Veranstaltung weiterentwickelt.
		Das Planspiel hilft mir die Inhalte meines Studienfachs besser zu verstehen.
HILVE		Ich lerne viel in der Veranstaltung.
		Mein Wissensstand ist nach der Veranstaltung wesentlich höher als vorher.
HILVE		Ich verfüge über ein grundlegendes Verständnis als vor der Veranstaltung.
		Ich lerne etwas Sinnvolles und Wichtiges.
Rahmenbedingungen		Die digitale Umgebung hat den Zweck der Veranstaltung gefördert.
		Es macht Sinn das Planspiel als Blockveranstaltung durchzuführen.
		Es wäre besser das Planspiel verteilt über das ganze Semester zu spielen.

		Der Zeitpunkt der Planspielveranstaltung ist sinnvoll in den Studienverlauf integriert.	
	HILVE	Das Thema der Veranstaltung als solches ist relevant (Beruf/Praxis/Prüfung/Gesellschaftlich).	
	HILVE	Das Tempo der Veranstaltung: (zu schnell bis zu langsam)	
		Die Anforderungen sind: (zu hoch bis zu niedrig)	
Planspiel		Das Planspiel ist ansprechend gestaltet.	
	Trautwein	Die zur Verfügung gestellten schriftlichen Informationen sind gut verständlich.	
		Die Berichte zu jeder Periode sind gut verständlich.	
		Die Spielergebnisse lassen sich nachvollziehen.	
		Die Spielergebnisse sind einleuchtend.	
		Die verfügbare Zeit je Spielrunde war...	
		Der Schwierigkeitsgrad des Planspiels war für mich...	
	Trautwein	Das Planspiel ist praxisnah.	
		Das Planspiel bildet die Realität gut ab.	
			Ich habe verstanden was ich tun muss, um im Planspiel erfolgreich zu agieren.
		Ich habe verstanden, wie das Planspiel funktioniert.	
Planspielleitung allgemein	Trautwein	Die Planspielleitung kennt sich mit dem Planspiel gut aus.	
		Die Veranstaltung war gut strukturiert.	
		Die Planspielleitung war engagiert.	
		Das Verhältnis von Planspielleitung und Studierenden war gut.	
	HILVE	Die Veranstaltung wird in interessanter Form gehalten.	
	HILVE	Die Veranstaltung zieht sich schleppend dahin.	
	Knödler	Die Planspielleitung setzte den Studierenden ein hohes Anforderungsniveau.	
Hühn	Die Planspielleitung war bei Fragen jederzeit ansprechbar.		
Debriefing/Interaktion	HILVE	Die Planspielleitung fördert Fragen und aktive Mitarbeit.	
		Diskussionen werden gut geleitet (Anregungen von Beiträgen, Eingehen auf Beiträge, Zeiteinteilung, Bremsen von Vielrednern).	
	Trautwein nach Fragebogen TU Dresden	Der Umfang der gemeinsamen Besprechung war...	
		Es bestand in ausreichendem Maß die Möglichkeit sich mit der Planspielleitung zu beraten.	
		Die Beratung findet in einer konstruktiven Atmosphäre statt.	
	Die Beratung ist hilfreich.		

	Knödler	Die Planspielleitung stellte komplexe Zusammenhänge dar, anstatt sich auf oberflächliches Faktenwissen zu beschränken.
		Die Planspielleitung verdeutlicht Inhalte mit Beispielen aus der Praxis.
	Hühn	Die Unterstützung der Planspielleitung war fachlich hilfreich.
Teamarbeit	Trautwein	Die Diskussionen in der Gruppe waren konstruktiv.
		Die Zusammenarbeit in der Gruppe war so, wie ich es mir vorstelle.
		Es fiel uns leicht, in der Gruppe Entscheidungen zu treffen.
		Wir haben in der Gruppe gut zusammengearbeitet.
		Wir sind zielgerichtet vorgegangen.
		Die zeitliche Abstimmung in der Gruppe hat gut funktioniert.
		Die Atmosphäre in der Gruppe war gut.
		Welchen Platz gemäß Aktienkurs/anderes Kriterium belegt unsere Gruppe derzeit
	Hühn	Ich würde sagen, dass unser Team sehr gut organisiert war.
	Trautwein	Jeder hatte eine Funktion in der Gruppe
Es gab eine klare Aufgabenverteilung in der Gruppe		
Ich wusste was meine Aufgaben sind.		
Studierende	(HILVE)	Ich beteilige mich mit Wortbeiträgen / bei Diskussionen
		Beim Einbringen der Beiträge fühle ich mich frei und äusserungsfähig (Falls keine Beiträge bitte frei lassen).
	Hühn	Wenn andere meine Rolle betrachten würden, würden sie mich zu den aktiveren in der Gruppe zählen.
		Wie häufig haben Sie vor dieser Veranstaltung schon an einem Planspiel teilgenommen?
		In den Auswertungsphasen (gemeinsame Besprechung der Spielergebnisse im Plenum) habe ich mich aktiv beteiligt.
		In den Entscheidungsphasen (Arbeit in der Kleingruppe) habe ich mich aktiv beteiligt.
		Ich spiele gerne Strategiespiele (z.B. Schach, Siedler von Catan)
	Trautwein	Mein betriebswirtschaftliches Wissen war zu Beginn der Veranstaltung: eher zu niedrig, genau richtig, eher zu hoch