

Lassen sich die Evaluationsergebnisse im Rahmen einer strukturierten Überarbeitung von Skripten einer Vorlesungsreihe verbessern?

Eine prospektive empirische Beobachtungsstudie

ABSTRACT/ZUSAMMENFASSUNG

Ziele: Die Studie zielte darauf ab die Frage zu beantworten, ob die strukturierte Überarbeitung von Vorlesungen im Vergleich zu nicht überarbeiteten Vorlesungen derselben Veranstaltungsreihe, ein signifikant besseres Evaluationsergebnis der Studierenden erzielt.

Methoden: Die vorliegende Studie untersuchte die Evaluationsergebnisse von Studierenden der Live Online-Lehrveranstaltung „Vorlesung Sportmedizin I (Orthopädie)“, im Wintersemester 2020/2021 an der Ruprecht-Karls-Universität Heidelberg. Als studienbedingte Maßnahme wurden die Vorlesungsmaterialien von 5 der 9 Veranstaltungen randomisiert ausgewählt und einer strukturierten Überarbeitung unterzogen, welche die Formulierung von Lernzielen, Gliederung in einzelne thematische Lernabschnitte und Wiederholungseinheiten am Ende der Vorlesung beinhaltete. Vier Vorlesungen wurden nicht überarbeitet (Vergleichsgruppe). Am Ende jeder Veranstaltung wurde eine anonyme Evaluation mit standardisierten, etablierten Fragenbögen durchgeführt, die an der Universität Heidelberg eingesetzt wurden. Die Studierenden waren über das Studienvorhaben nicht explizit informiert.

Ergebnisse: Insgesamt wurden 300 vollständige Evaluationsbögen für 9 Veranstaltungen ausgewertet (strukturierte Überarbeitung: $n=143$, Vergleichsgruppe ohne Überarbeitung: $n=157$ Datensätze). Zwischen überarbeiteten und nicht überarbeiteten Vorlesungen zeigte sich kein statistisch signifikanter Unterschied der Evaluationsergebnisse mit Hinblick auf die Definition von Lernzielen, der Strukturierung der Vorlesungseinheit, dem Eingehen auf Fragen und Belange der Studierenden durch den Dozierenden sowie den selbst eingeschätzten Lernzuwachs ($p>0,05$; Kruskal-Wallis-Tests).

Schlussfolgerung: Im Rahmen der studienbedingten Maßnahme zeigte sich kein signifikant besseres Evaluationsergebnis der Studierenden.

Level of Evidence: Level II, prospektive, Beobachtungsstudie mit Vergleichsgruppe.

Schlagworte: Vorlesung – Evaluation – Struktur – Lernziel – Selbsteinschätzung Lernzuwachs

Aims: This study aimed to determine whether a structured revision of lecture course scripts would be associated with better student evaluation results.

Methods: The present study investigated the evaluation results of the live online lecture course “Sports Medicine I (Orthopaedics)”, during fall semester 2020/2021 at the Ruprecht-Karls-University Heidelberg, Germany. Five of the nine lectures were ran-

domly chosen and a structured revision of lecture course scripts was performed (definition of learning aims, arrangement in small learning chapters and a repetition at the end of each lecture with questions and answers). Four lectures were not revised (control-group). At the end of each lecture, an anonymous evaluation was performed by the students using an established, standardized questionnaire. The students were not informed about the study.

Results: In total, 300 questionnaires were available for analysis (revised lectures group: n=143, non-revised lectures group: n=157). No significant difference could be found for the evaluation results between lectures with revised and unrevised lecture scripts, regarding the definition of learning aims, structure of each lecture, giving the opportunity for questions and answers and the self-ranked increase in learning ($p>0,05$; Kruskal-Wallis-Tests).

Conclusion: No improvement in evaluation results could be found after a structured revision of lecture course scripts.

Level of Evidence: Level II, prospective, randomized, controlled, observational study

Keywords: Lecture – Evaluation – Structure – Self-evaluation of increase in learning

Einleitung

Die Vorlesungsreihe „Sportmedizin I (Orthopädie)“ ist seit mehreren Jahren etablierter Bestandteil des Bachelor-Studiengangs Sportwissenschaft (B. Sc.) am Institut für Sport und Sportwissenschaft (ISSW) der Ruprecht-Karls-Universität Heidelberg. Ziel der Vorlesungsreihe ist den Studierenden grundlegende Kenntnisse im Bereich Anatomie und Orthopädie zu vermitteln. Die übergeordneten Lehrziele der Vorlesungsreihe sind Wissen über die Funktion des Bewegungsapparats, potentielle Verletzungen oder Überlastungsreaktionen und deren Entstehungsmechanismen bzw. vorbeugende Maßnahmen (Prävention) und zuletzt auch im Falle einer aufgetretenen Verletzung über Erstmaßnahmen, Therapiemöglichkeiten und potentielle Langzeitauswirkungen zu vermitteln.

Zu Beginn der Vorlesungsreihe, vor mehreren Jahren, wurden für diesen Zweck die bereits vorhanden Vorlesungsslides für Medizinstudierende eingesetzt, wobei der Umfang der Slides leicht gekürzt und Inhalte vereinfacht wurden. Jedoch zeigte sich im Verlauf der Jahre, dass die Gliederung und der Inhalt der bestehenden Slides unzureichend geeignet war, um die o.g. Lehrzieleerreichung im Bachelor-Studiengang Sportwissenschaft leicht verständlich zu ermöglichen. Insbesondere waren in den bisherigen Vorlesungsunterlagen die Lernziele unzureichend definiert und daher auch in den jeweiligen einzelnen Vorlesungen für die Studierenden kaum ersichtlich. Auf Grund der Struktur und des Inhalts der Vorlesungsunterlagen ergab sich tendenziell ein stark lehrendenzentrierter Lehransatz, wobei die Lehrperson die aktive Rolle und die Studierenden die passive Rolle einnahmen.

Eine wichtige Voraussetzung für langfristiges erfolgreiches Lernen ist, den Lernenden in einer aktiven Rolle einzubinden. Voraussetzung hierfür ist die Formulierung von Lernzielen und Förderung der aktiven Rolle der Studierenden. Daher zielte die aktuelle Studie

darauf ab die Frage zu beantworten, ob die Überarbeitung von Vorlesungsveranstaltungen, mittels Formulierung von Lernzielen, Gliederung in einzelne thematische Lernschritte und Wiederholungseinheiten am Ende der Vorlesung, im Vergleich zu nicht überarbeiteten Vorlesungen derselben Veranstaltungsreihe, ein signifikant besseres Evaluationsergebnis der Studierenden erzielt.

Methoden

Studiendesign

Die vorliegende, prospektive, empirische Beobachtungsstudie mit Vergleichsgruppe untersuchte die anonymen Evaluationsergebnisse von Studierenden der Lehrveranstaltung „Vorlesung Sportmedizin I (Orthopädie)“, im Wintersemester 2020/2021 an der Ruprecht-Karls-Universität Heidelberg (Veranstaltungsnummer im Vorlesungsverzeichnis: 10SPO40013). Die anonymen Evaluationsergebnisse wurden vom Studiendekanat der Ruprecht-Karls-Universität Heidelberg zur Verfügung gestellt. Auf Grund der Auswertung von anonymen Daten wurde kein Ethikvotum oder Einverständniserklärung der Studierenden benötigt. Das Studiendesign und die methodische Auswertung der Daten wurde bereits vor Studienbeginn festgelegt. Der Autor dieser Studie ist deckungsgleich mit der Lehrperson der Veranstaltungsreihe.

Die Vorlesungsreihe „Sportmedizin I (Orthopädie)“ findet nur im Wintersemester statt und hat einen Umfang von 11 Veranstaltungen je zwei Semesterwochenstunden und zusätzlichem Klausurtermin am Ende der Vorlesungsreihe. Für die meisten Studierenden ist die Teilnahme an der Veranstaltung verpflichtend, wobei jedoch jährlich ca. 10% der Teilnehmenden die Veranstaltung als Wahlfach wählen. Ein Termin der Vorlesungsreihe wird fachfremd aus der Inneren Medizin über das Thema Doping unterrichtet und an einem weiteren Termin wird eine Wiederholungseinheit vor der Klausur durchgeführt. Somit werden in der Vorlesungsreihe „Sportmedizin I (Orthopädie)“ 9 Veranstaltungen abgehalten, die sich auf die Vermittlung von neuem orthopädischem bzw. unfallchirurgischem Wissen fokussieren (Tabelle 1). Die vorliegende Studie untersucht die Daten dieser 9 Veranstaltungen.

Auf Grund der COVID-19 Pandemie wurde die „Vorlesung Sportmedizin I (Orthopädie)“ im Wintersemester 2020/2021 erstmals online als Live-Veranstaltung über die elektronische Plattform „heiCONF Audimax“ durchgeführt. Die Studierenden können nach einer Registrierung an der Live-Veranstaltung teilnehmen, wobei Fragen entweder über eine Chat-Funktion oder per Mikrofon gestellt werden können. Die Vorlesungsmaterialien (Powerpoint-Slides) können durch den Dozierenden auf die Plattform geladen und präsentiert werden, wobei die Lehrperson live die Slides kommentieren und ggf. auch Zeichnungen oder Highlights setzen kann. Die Vorlesungsmaterialien (Powerpoint-Slides) der gesamten Veranstaltungsreihe zeigten bis zum vorangegangenen Wintersemester 2019/2020 eine

einheitliche Struktur, wobei die Veranstaltung bis einschließlich des Wintersemesters 2019/2020 im Sinn einer „klassischen Vorlesung“ in einem Hörsaal abgehalten wurde.

Studienbedingte Maßnahme

Im Rahmen der vorliegenden Studie wurden die Vorlesungsmaterialien (Powerpoint-Slides) von 5 der 9 Veranstaltungen aus dem Wintersemester 2019/2020 randomisiert ausgewählt (<https://www.randomizer.org>) und für das Wintersemester 2020/2021 einer strukturierten Überarbeitung unterzogen (Tabelle 1).

Datum	Veranstaltungstitel	Teilnehmer (n)	Vollständige Evaluationsbögen (n)	Prozentsatz der Evaluationen
5.11.2020	Einführung und Funktionelle Anatomie I	89	63	71
12.11.2020	Funktionelle Anatomie II	85	35	41
19.11.2020 #	Frakturlehre	87	27	31
26.11.2020 #	Kindertraumatologie	83	32	39
3.12.2020 #	Wirbelsäule	86	26	30
10.12.2020 #	Sporttraumatologie I: Obere Extremität	82	35	43
17.12.2020	Sporttraumatologie II: Hüfte und Knie	84	31	37
14.1.2021*	Doping	80	3	4
21.1.2021 #	Sporttraumatologie III: Sprunggelenk und Fuß	84	23	27
28.1.2021	Arthrose und rheumatische Erkrankungen	85	28	33
4.2.2021*	Wiederholung	85	1	1

Tabelle 1

Studienkohorte der Vorlesungsreihe Sportorthopädie I WS 2020/2021. * Diese Veranstaltungen wurden aus der Auswertung ausgeschlossen. # Diese Vorlesungen wurden im Sinn der Studienbezogenen Maßnahme strukturell überarbeitet.

Die Vorlesungsmaterialien der übrigen 4 Vorlesungen blieben unverändert im Vergleich zu den Jahren zuvor (Vergleichsgruppe). Die strukturierte Überarbeitung der Slides umfasste jeweils die Gliederung in drei bis vier Themenblöcke, Darstellung der Gliederung am Anfang jeder Präsentation, und Formulierung der Lernziele für jeden Abschnitt der neu ge-

gliederten Veranstaltung, Einfügen von Frage & Antwort-Slides nach jedem neu gegliederten Abschnitt um den Fluss der Vorlesung kurz zu unterbrechen und Fragen der Studierenden zu beantworten, sowie einer Wiederholungseinheit am Ende der Vorlesung. Die strukturierte Überarbeitung ist schematisch in Tabelle 2 dargestellt.

Vorlesungsstruktur vor der Überarbeitung	Vorlesungsstruktur nach der Überarbeitung
Titel (Slide 1)	Titel (Slide 1)
Organisatorisches (Slide 2)	Organisatorisches (Slide 2)
Themenblock 1-3 (ca. 40 Slides)	Wiederholung der Inhalte der letzten Vorlesung (ca. 3 - 6 Slides)
Zusammenfassung (1 Slide)	Lernziele und Gliederung der heutigen Vorlesung (eine Slide)
	Themenblock 1 (ca. 10 Slides)
	Frage & Antwort Themenblock 1 (1 Slide)
	Lernziele und Gliederung der heutigen Vorlesung (eine Slide)
	Themenblock 2 (ca. 10 Slides)
	Frage & Antwort Themenblock 1 (2 Slide)
	Lernziele und Gliederung der heutigen Vorlesung (eine Slide)
	Themenblock 3 (ca. 10 Slides)
	Frage & Antwort Themenblock 3 (2 Slide)
	Zusammenfassung an Hand der Lernziele und Erinnerung an die Evaluation (1 Slide)

Tabelle 2

Schematische Darstellung der Überarbeitung der Vorlesungsmaterialien.

Datenerhebung

Zu Beginn der ersten Veranstaltung wurden die Studierenden informiert, dass die Evaluation der Lehrveranstaltung nicht nur wie sonst üblich, einmalig am Ende der Veranstaltungsreihe erfolgt, sondern am Ende jeder Veranstaltung um ein genaueres Bild jeder einzelnen Vorlesung zu erhalten. Hierfür wurde auf der jeweils letzten Slide jeder Vorlesungspräsentation der Internet-Link sowie ein QR-Code mit dem eingebetteten Internet-Link zur Evaluation der jeweiligen Veranstaltung einblendet. Die Studierenden wurden am Ende der Veranstaltungen (n=9) daran erinnert, an der Evaluation teilzunehmen um eine statistische Auswertung mit validen Ergebnissen zu ermöglichen, mit der Absicht die Vorlesungsmaterialien mittelfristig zu verbessern. Die Studierenden wurden nicht explizit über das Studienvorhaben informiert. Die Fragen der Evaluation orientierten sich an den bisher verwendeten, standardisierten Fragebögen, die zur Evaluation an der Universität Heidelberg eingesetzt werden. Wie in der statistischen Literatur empfohlen, wird jede Frage auf einer Likert Skala mit mindestens 5-Punkten beantwortet, die von „stimme voll zu“ bis „stimme nicht zu“ reicht (HINKIN, TRACEY & ENZ 1997). Für die Fragen 1.2) bis 2.9) stellen niedrigere im Vergleich zu höheren Punktwerten auf der Likert Skala bessere Evaluationsergebnisse dar. Die Frage 2.10) wird auf einer 10-Punkte Likert Skala beantwortet und dient der Erstellung des international, vor allem im Bereich der Konsumforschung verwendeten Net-Promotor-Scores (NPS), der 2003 erstmal von F. Reichheld im Harvard Business Review beschrieben wurde (REICHHELD 2003). Zusammenfassend ist der NPS eine individuell berechnete Kennzahl, die darüber Auskunft gibt wie wahrscheinlich ein Konsument ein Produkt oder eine Dienstleistung weiterempfehlen würde, wobei ein Score von -100 den schlechtesten und +100 den besten Wert darstellt (REICHHELD 2003). Um die Datenerhebung pro Evaluation auf jeweils nur eine Veranstaltung einzugrenzen, wurden die Fragen leicht modifiziert indem als Überschrift „Bitte bewerten Sie die heutige Veranstaltung“ ergänzt wurde, woraufhin die einzelnen Evaluationsfragen folgten. Der Evaluationsbogen ist in Abbildung 1 (umseitig) dargestellt.

Evaluation einer strukturierten Überarbeitung von Skripten einer Vorlesungsreihe

UNIVERSITÄT
HEIDELBERG

Dozenten und Einzelveranstaltungen WS20/21

Medizinische Fakultät Heidelberg
 Innmann, Sportorthopädie

Kontrastmodus aktivieren

0 %

Vorlesung Sportorthopädie (PD Dr. M. Innmann)

1 Allgemeine Fragen

1.1 Heutiges Datum:*	<input type="radio"/>	5.11.2020	<input type="radio"/>	12.11.2020	<input type="radio"/>	19.11.2020	<input type="radio"/>	26.11.2020	<input type="radio"/>	3.12.2020
	<input type="radio"/>	10.12.2020	<input type="radio"/>	17.12.2020	<input type="radio"/>	7.1.2021	<input type="radio"/>	14.1.2021	<input type="radio"/>	21.1.2021
	<input type="radio"/>	28.1.2021	<input type="radio"/>	4.2.2021	<input type="radio"/>	11.2.2021	<input type="radio"/>	18.2.2021	<input type="radio"/>	25.2.2021

sehr groß
sehr gering

1.2 Mein Interesse am Fach vor der Veranstaltung war	<input type="radio"/>					
1.3 Mein Interesse am Fach nach der Veranstaltung ist	<input type="radio"/>					

2 Bitte bewerten Sie die heutige Veranstaltung

	stimme voll zu		stimme nicht zu
2.1 Die Lernziele wurden zu Beginn klar definiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.2 Die Lehrveranstaltung ist klar strukturiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.3 Der Lehrstoff wird in verständlicher Weise vermittelt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.4 Es wird auf Fragen und Belange der Studierenden eingegangen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.5 Mein Lernzuwachs ist hoch.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.6 Der Dozent ist engagiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.7 Der Dozent bezieht mich aktiv in die Lehre ein (durch Fragen, Diskussionen, Aufgabenstellungen etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.8 Die Lernatmosphäre ist angenehm.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

außergewöhnlich gut
außergewöhnlich schlecht

2.9 Die Veranstaltung insgesamt finde ich	<input type="radio"/>					
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

sehr unwahrscheinlich
sehr wahrscheinlich

2.10 Mit welcher Wahrscheinlichkeit würden Sie diese Veranstaltung an befreundete Studierende des gleichen Studiengangs weiterempfehlen?	<input type="radio"/>								
--	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Absenden

Abbildung 1

Darstellung des Online-Evaluationsbogens, der von den Studierenden am Ende jeder einzelnen Vorlesungsveranstaltung ausgefüllt werden konnte.

Insgesamt lagen 304 vollständige Evaluationsbögen für alle 11 Veranstaltungen vor, wobei nach Ausschluss der Fragebögen (n=4) von den oben genannten 2 Veranstaltungen, mit nicht orthopädischem bzw. unfallchirurgischem Inhalt am 14.01.2021 und 04.02.2021, 300 vollständige Evaluationsbögen vorlagen, die die Datengrundlage dieser Studie bilden und näher untersucht werden sollen. Die Mindestzahl an vollständigen Evaluationsbögen pro Veranstaltung lag bei 23 (vergleiche Abbildung 2A&B).

Statistische Auswertung

Die anonymisierten Daten der Evaluation wurden vom Dekanat der Universität Heidelberg in Form einer bereits ausgewerteten Evaluation (Abbildung 2A&B) sowie einer Excel-Tabelle mit Rohdaten zur Verfügung gestellt, wobei die Daten der Evaluationsbögen in der Tabelle zu den jeweiligen Vorlesungsveranstaltungen zugeordnet werden konnten. Im Anschluss wurden die Daten in eine SPSS-Datei übertragen und die einzelnen Evaluationsbögen danach codiert, ob sie für eine Veranstaltung mit strukturierter Überarbeitung oder ohne strukturierte Überarbeitung ausgefüllt wurden. Nach dieser Aufteilung lagen für überarbeitete Vorlesungen 143 und für nicht überarbeitete Vorlesungen 157 ausgefüllte Evaluationsbögen vor. Als erster Schritt der statistischen Auswertung erfolgte eine explorative Datenanalyse und Tests auf Normalverteilung. Es zeigte sich hierbei keine Normalverteilung der Daten, weshalb im Anschluss nicht-parametrische Tests verwendet wurden. Kruskal-Wallis Tests wurden durchgeführt um die Fragestellung der Studie zu beantworten. Im Detail wurden die Ergebnisse für jede einzelne Evaluationsfrage mit einem separaten Kruskal-Wallis Test ausgewertet, wobei die Daten für jede einzelne Evaluationsfrage über alle Vorlesungen aggregiert wurden. Als Gruppierungsvariable wurde ausgewählt, ob die Evaluationsergebnisse für eine überarbeitete oder nicht überarbeitete Vorlesung abgegeben wurden. Paarweise Vergleiche wurden mit Wilcoxon-Signed-Ranks Tests durchgeführt. Der Net-Promotor-Score wurde für das Gesamtkollektiv ($n=300$), sowie die Untergruppen mit überarbeiteten und nicht überarbeiteten Vorlesungsmaterialien berechnet (<http://www.npscalculator.com/de> Last Access: 12.03.2021; 10:55 Uhr). Ob ein statistisch signifikanter Unterschied zwischen den beiden zuletzt genannten Untergruppen für die Verteilung an Promotoren (Score auf der Likert Skala ≥ 9) und Distraktoren (Score auf der Likert Skala ≤ 6) vorliegt, wurde mittels eines Chi²-Test berechnet. Bei p-Werten von 0,05 oder weniger wurde angenommen, dass ein statistisch signifikanter Unterschied vorliegt. Sofern nicht anders gekennzeichnet, werden die Daten als Mittelwert (MW) mit Standardabweichung angegeben (SD). Um eine ausreichende statistische Aussagekraft der Auswertung zu erreichen, wurde die benötigte Stichprobengröße, auf Basis bisheriger Werte in der Literatur, auf mindestens 20 Messwerte (ausgefüllte Evaluationsbögen) pro Veranstaltung und mindestens 300 Messwerten insgesamt abgeschätzt (SAMUELS 2015). Die statistische Auswertung erfolgte mit dem Programm SPSS Version 27.0 (IBM SPSS Statistics, IBM, Armonk, NY, USA).

Ergebnisse

Evaluationsergebnis der gesamten Vorlesungsreihe

Für die gesamte Vorlesungsreihe, d.h. ohne getrennte Betrachtung ob überarbeitete oder nicht überarbeitete Vorlesungen gehalten wurden, zeigte sich ein großes Interesse der Studierenden am Fach direkt vor (Frage 1.2: $1,89 \pm 0,8$) und direkt nach der jeweiligen Veranstaltung (Frage 1.3: $1,74 \pm 0,79$) (vgl. Tabelle 3). Dabei war das Interesse der Studierenden am Fach direkt nach der jeweiligen Veranstaltung mit hoher statistischer Signifikanz geringfügig höher ausgeprägt als vor der Veranstaltung ($p < 0,001$; Wilcoxon-Signed-Ranks Test).

Ansonsten zeigten sich für die gesamte Veranstaltung eine hohe Zustimmung der Studierenden dafür, dass die Lernziele klar definiert wurden (Frage 2.1: $1,45 \pm 0,72$); die Lehrveranstaltung klar strukturiert war (Frage 2.2: $1,44 \pm 0,69$), der Lehrstoff in verständlicher Weise vermittelt wurde (Frage 2.3: $1,44 \pm 0,69$), auf Fragen und Belange der Studierenden eingegangen wurde (Frage 2.4: $1,19 \pm 0,53$), der Lernzuwachs hoch war (Frage 2.5: $1,8 \pm 0,87$), die Lehrperson engagiert war (Frage 2.6: $1,25 \pm 0,59$) und die Lernatmosphäre als angenehm empfunden wurde (Frage 2.8: $1,71 \pm 0,8$). Allerdings zeigte sich auch, dass sich die Studierenden nur „durchschnittlich“ aktiv in die Lehre einbezogen fühlten (Frage 2.7: $2,59 \pm 1,14$). Insgesamt wurde die Veranstaltung als gut eingeschätzt (Frage 2.9: $1,99 \pm 0,82$; vergleiche Abbildung 2A & B).

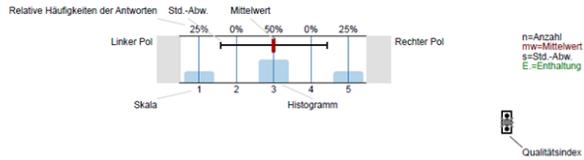
Die Studierenden schätzten die Wahrscheinlichkeit als hoch ein, die Veranstaltung an befreundete Studierende des gleichen Studiengangs weiterzuempfehlen (Frage 2.10: $8,27 \pm 1,82$). Der Net-Promotor-Score für das Gesamtkollektiv betrug 38 Punkte.

Innmann, Sportorthopädie
Erfasste Fragebögen n=305



Legende

Frage**text**

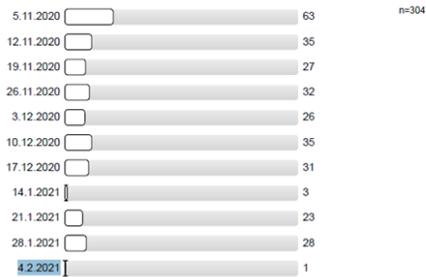


Erklärung der Ampelsymbole

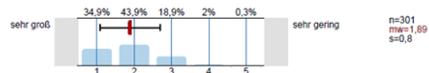
- Der Mittelwert liegt unterhalb der Qualitätsrichtlinie.
- Der Mittelwert liegt im Toleranzbereich der Qualitätsrichtlinie.
- Der Mittelwert liegt innerhalb der Qualitätsrichtlinie.

1. Allgemeine Fragen

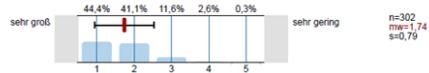
1.1) Heutiges Datum:



1.2) Mein Interesse am Fach vor der Veranstaltung war

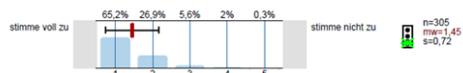


1.3) Mein Interesse am Fach nach der Veranstaltung ist



2. Bitte bewerten Sie die heutige Veranstaltung

2.1) Die Lernziele wurden zu Beginn klar definiert.



2.2) Die Lehrveranstaltung ist klar strukturiert.

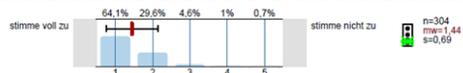


Abbildung 2A

Auswertung der Evaluation vom Dekanat der Universität Heidelberg (Seite 1).

Evaluation einer strukturierten Überarbeitung von Skripten einer Vorlesungsreihe

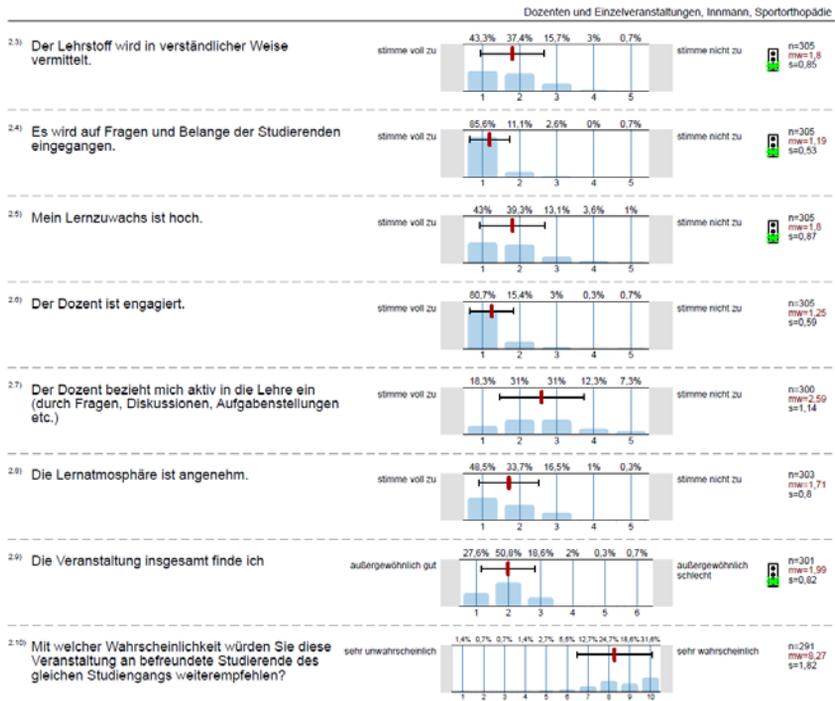


Abbildung 2B

Auswertung der Evaluation vom Dekanat der Universität Heidelberg (Seite 2).

Evaluationsergebnis mit getrennter Betrachtung für überarbeitete und nicht überarbeitete Vorlesungen

Bei der getrennten Betrachtung der Evaluationsergebnisse für überarbeitete und nicht überarbeitete Vorlesungen zeigte sich, dass ein hohes Interesse der Studierenden am Fach direkt vor und direkt nach der jeweiligen Veranstaltung vorlag, jedoch bei überarbeiteten Vorlesungen etwas niedriger als bei nicht überarbeiteten Vorlesungen war (Frage 1.2: $1,96 \pm 0,78$ vs. $1,77 \pm 0,76$; $p=0,04$ und Frage 1.3: $1,86 \pm 0,86$ vs. $1,58 \pm 0,69$; $p=0,01$; Kruskal-Wallis-Tests).

Die Evaluation zeigte zwischen überarbeiteten und nicht überarbeiteten Vorlesungen keine signifikanten Unterschiede hinsichtlich der Definition von Lernzielen (Frage 2.1: $1,37 \pm 0,61$ vs. $1,47 \pm 0,73$; $p=0,28$), Strukturierung der Lehrveranstaltungen (Frage 2.2: $1,37 \pm 0,65$ vs. $1,46 \pm 0,643$; $p=0,24$), der verständlichen Vermittlung des Lehrstoffs (Frage 2.3: $1,72 \pm 0,79$ vs. $1,82 \pm 0,85$; $p=0,14$), des Eingehens auf Fragen und Belange der Studierenden (Frage 2.4: $1,19 \pm 0,45$ vs. $1,12 \pm 0,37$; $p=0,10$), des Lernzuwachses (Frage 2.5: $1,8 \pm 0,84$ vs. $1,73 \pm 0,85$; $p=0,38$), des Engagement der Lehrperson (Frage 2.6: $1,28 \pm 0,61$ vs. $1,18 \pm 0,44$; $p=0,08$), der aktiven Einbeziehung der Studierenden in die Lehre (Frage 2.7: $2,48 \pm 1,07$ vs. $2,68 \pm 1,23$; $p=0,36$), der Lernatmosphäre (Frage 2.8: $1,75 \pm 0,77$ vs. $1,63 \pm 0,75$; $p=0,14$), der allgemeinen Einschätzung der Veranstaltung (Frage 2.9: $5,07 \pm 0,82$ vs. $5,04 \pm 0,76$; $p=0,54$) und der Wahrscheinlichkeit die Veranstaltung an befreundete Studierende des gleichen Studiengangs weiterzuempfehlen (Frage 2.10: $8,11 \pm 1,95$ vs. $8,52 \pm 1,56$; $p=0,19$; Kruskal-Wallis-Tests), zeigte (vergleiche Tabelle 3 und Abbildung 3).

Der Net-Promotor-Score für Veranstaltungen mit überarbeiteten und nicht überarbeiteten Vorlesungen betrug 43 bzw. 33 Punkte, jedoch ohne statistisch signifikanten Unterschied in der Verteilung der Promotoren bzw. Distraktoren ($p=0,29$; Chi2-Test; vergleiche Tabelle 4).

Evaluationsfrage	Vorlesung überarbeitet (Mittelwert ± SD)	Vorlesung nicht überarbeitet (Mittelwert ± SD)	<i>p</i> -Wert
1.2: Mein Interesse am Fach vor der Veranstaltung war	1,96 ± 0,78	1,77 ± 0,76;	0,04
1.3: Mein Interesse am Fach nach der Veranstaltung ist	1,86 ± 0,86	1,58 ± 0,69	0,01
2.1: Die Lernziele wurden zu Beginn klar definiert	1,37 ± 0,61	1,47 ± 0,73	0,28
2.2: Die Lehrveranstaltung ist klar strukturiert	1,37 ± 0,65	1,46 ± 0,64	0,24
2.3: Der Lehrstoff wird in verständlicher Weise vermittelt	1,72 ± 0,79	1,82 ± 0,85	0,14
2.4: Es wird auf Fragen und Belange der Studierenden eingegangen	1,19 ± 0,45	1,12 ± 0,37	0,10
2.5: Mein Lernzuwachs ist hoch	1,8 ± 0,84	1,73 ± 0,85	0,38
2.6: Der Dozent ist engagiert	1,28 ± 0,61	1,18 ± 0,44	0,08
2.7: Der Dozent bezieht mich aktiv in die Lehre ein	2,48 ± 1,07	2,68 ± 1,23	0,36
2.8: Die Lernatmosphäre ist angenehm	1,75 ± 0,77	1,63 ± 0,75	0,14
2.9: Die Veranstaltung insgesamt finde ich	5,07 ± 0,82	5,04 ± 0,76	0,54
2.10: Mit welcher Wahrscheinlichkeit würden Sie diese Veranstaltung weiterempfehlen?	8,11 ± 1,95	8,52 ± 1,56	0,19

Tabelle 3

Darstellung der Evaluationsergebnisse für die Fragen 1.2) bis 2.10) getrennt für überarbeitete und nicht überarbeitete Vorlesungen (Kruskal-Wallis-Tests).

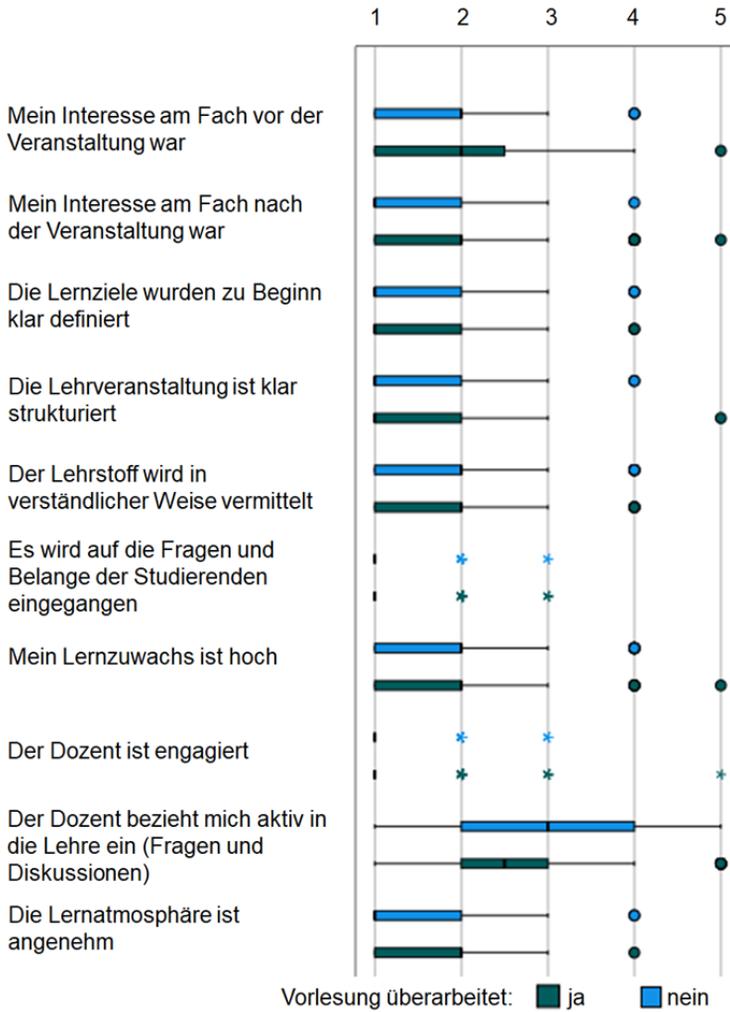


Abbildung 3

Boxplots zur Darstellung der Evaluationsergebnisse für die Fragen 1.2) bis 2.9) getrennt für überarbeitete (dunkelgrün) und nicht überarbeitete Vorlesungen (hellblau).

	Detraktor	Promotor	Gesamt
Vorlesung überarbeitet	25	64	89
Vorlesung nicht überarbeitet	22	80	102
Gesamt	47	144	191

Tabelle 4

Verteilung der Promotoren (Score auf der Likert Skala ≥ 9) und Distraktoren (Score auf der Likert Skala ≤ 6) für überarbeitete und nicht überarbeitete Vorlesungen. Die übrigen 109 Evaluationsbögen zeigten indifferente Scores auf der Likert Skala zwischen 7 und 8 und sind in dieser Tabelle nicht aufgelistet ($p=0,29$; Chi-2-Test).

Diskussion

Die vorliegende, prospektive, empirische Beobachtungsstudie konnte keinen relevanten Unterschied der Evaluationsergebnisse zwischen Vorlesungsveranstaltungen mit strukturiert überarbeiteten Vorlesungsunterlagen und Vorlesungsveranstaltungen ohne Überarbeitung zeigen. In beiden Gruppen lagen vergleichbar gute Evaluationsergebnisse vor, wobei sich allerdings auch zeigte, dass sich die Studierenden nur durchschnittlich aktiv in die Lehre einbezogen fühlten. Die prozentuale Teilnahme an den Evaluationen lag auf einem akzeptablen Niveau zwischen 27% und 71%, wobei die notwendige Mindestzahl an kompletten Evaluationsbögen in Höhe von 20 Stück pro Vorlesung nicht unterschritten und die Gesamtzahl von benötigten Datensätzen ($n=300$) erreicht wurde. Somit liegt eine ausreichende statistische Aussagekraft der erhobenen Daten vor, die eine sinnvolle Interpretation der Ergebnisse ermöglicht.

Die aktuellen Daten dieser Studie sollen in drei thematischen Abschnitten interpretieren werden. Erstens muss sichergestellt werden, dass keine relevanten Störfaktoren die Auswertung der Evaluationsergebnisse hinsichtlich der studienbedingten Maßnahme beeinflussen. Da das Interesse der Studierenden am Fach vor der Veranstaltung und auch das Engagements der Lehrperson potentielle Einflussfaktoren auf das Evaluationsergebnis und auf den Lernzuwachs der Studierenden sein könnten, muss ein relevanter Unterschied dieser Faktoren, im Sinn von Störfaktoren hinsichtlich der studienbedingten Maßnahme ausgeschlossen werden. Bei getrennter Betrachtung für überarbeitete und nicht überarbeitete Vorlesungen zeigte sich nur eine sehr geringe Differenz für das Interesse der Studierenden am Fach vor der jeweiligen Veranstaltung (Frage 1.2: $1,96 \pm 0,78$ vs. $1,77 \pm 0,76$; $p=0,04$). Obwohl der Unterschied statistisch signifikant war, muss diese Differenz zu Gunsten der nicht überarbeiteten Vorlesungen als wenig relevant eingeschätzt werden, da das Ausmaß

der Differenz zu gering ausgeprägt war (Effektstärke 0,25). Somit darf angenommen werden, dass das Interesse der Studierenden am Fach vor den jeweiligen Veranstaltungen beider Gruppen im Wesentlichen gleich war. Des Weiteren zeigte sich kein Unterschied der Evaluationsergebnisse hinsichtlich des Engagements der Lehrperson zwischen überarbeiteten und nicht überarbeiteten Vorlesungen (Frage 2.6: $1,28 \pm 0,61$ vs. $1,18 \pm 0,44$; $p=0,08$). Somit ist der Vergleich der Evaluationsergebnisse der beiden Gruppen unter Berücksichtigung der studienbedingte Maßnahme erlaubt. Des Weiteren wurden die Studierenden nicht explizit über das Studienvorhaben informiert. Insofern ist ausgeschlossen, dass eine Beeinflussung der Studierenden vorlag, auf Grund der Kenntnis des Studienvorhabens, die überarbeiteten Vorlesungen besser oder schlechter zu evaluieren.

Zweitens sollte die Interpretation der Evaluationsergebnisse hinsichtlich der objektiven studienbedingten Maßnahmenkriterien erfolgen. Wesentlicher Bestandteil der studienbedingten Maßnahme war die klare Formulierung von Lernzielen und Strukturierung bzw. Gliederung der Vorlesungen. Bemerkenswerterweise zeigte sich in der Evaluation für genau diese Punkte, dass die Studierenden keinen Unterschied zwischen überarbeiteten und nicht überarbeiteten Vorlesungen angaben (Frage 2.1: $1,37 \pm 0,61$ vs. $1,47 \pm 0,73$; $p=0,28$ und 2.2: $1,37 \pm 0,65$ vs. $1,46 \pm 0,643$; $p=0,24$). Ebenso wenig zeigte sich kein Unterschied in der Evaluation, inwiefern auf Fragen und Belange der Studierenden eingegangen wurde, obwohl in den überarbeiteten Vorlesungen Frage & Antwort-Slides nach jedem neu gegliederten Abschnitt integriert wurden um Fragen der Studierenden zu beantworten, was in den Vorlesungen von Seiten der Studierenden auch angenommen wurde (Frage 2.4: $1,19 \pm 0,45$ vs. $1,12 \pm 0,37$; $p=0,10$). Da die Evaluation lediglich nach Werten auf der Likert-Skala fragt und keine Begründung im Sinn eines Freitextes gegeben werden konnte, kann die aktuelle Studie letztlich nicht beantworten, warum für diese studienbedingten Maßnahmenkriterien kein Unterschied der Evaluationsergebnisse festgestellt werden konnte.

Drittens sollte die Interpretation der Evaluationsergebnisse hinsichtlich der subjektiven Einschätzung des Lernzuwachses der Studierenden erfolgen. Hierbei zeigte sich ebenfalls, dass die Studierenden im Durchschnitt zustimmten, einen hohen Lernzuwachs zu haben, wobei jedoch erneut kein Unterschied zwischen überarbeiteten und nicht überarbeiteten Vorlesungen festgestellt werden konnte (Frage 2.5: $1,8 \pm 0,84$ vs. $1,73 \pm 0,85$; $p=0,38$). Somit scheint die studienbedingte Maßnahme, keinen signifikanten Einfluss auf den selbst eingeschätzten Lernzuwachs der Studierenden gehabt zu haben. Diese Folgerung sollte in Zusammenschau mit dem Evaluationsergebnis von Frage 2.7 betrachtet werden, die danach fragte, inwieweit sich die Studierenden aktiv in die Lehre einbezogen fühlen. Bei der Beantwortung dieser Frage zeigte sich bemerkenswerterweise, dass die Studierenden für beide Gruppen der Vorlesungen nur mittelmäßig zustimmten, aktiv in die Lehre einbezogen zu werden. Insbesondere zeigte sich für diese Frage im Vergleich zu den Fragen 1.2 bis 2.8 ein „schlechteres“ Evaluationsergebnis, was möglicherweise am grundsätzlichen Design des Lehrformats „Vorlesung“ liegen mag, jedoch sicherlich auch am lehrenden- und nicht lernenden-zentrierten Lehransatz.

Die aktuellen Ergebnisse lassen sich im Kontext der Literatur wie folgt interpretieren. In der Literatur zur Lerntheorie zeigt sich, dass nach Marton und Säljö im Wesentlichen zwei gegenpolige Ansätze des Lernens existieren. Einerseits das oberflächliche (defensive) Lernen und andererseits das tief reichende (expansive) Lernen. Wesentliche Voraussetzung des Letzteren ist, dass eine Aktivierung im Sinn einer aktiven Teilnahme des Lernenden an diesem Prozess erforderlich ist (MARTON & SÄLJÖ 1976 a + b). Dies zeigt sich ebenfalls in der Lernforschung adaptierten (therapeutischen) Ansatz der Aktiven Partnerschaft von Steve de Shazer (DE SHAZER 1994), die letztlich darauf abzielt, die Lernenden als aktive und selbstverantwortliche Kooperationspartner*innen im Lernprozess zu gewinnen. Dies wird im Wesentlichen dadurch ermöglicht, dass Lehrende die Verantwortung für folgende drei Punkte übernehmen: Eine Lernvereinbarung mit den Lernenden zu schließen, die Rolle von Anleiter*innen / Unterstützer*innen („Facilitator“) einnehmen und aktivierende Lernmethoden anwenden bzw. aktivierende Lernsituationen schaffen.

Wenn die Struktur und der Inhalt der aktuellen Vorlesungsreihe sowie die studienbedingte Maßnahme mit Hinblick auf die o.g. lerntheoretischen Aspekte in der Literatur analysiert werden, so muss kritisch festgestellt werden, dass im Wesentlichen weiterhin das Design einer lehrendenzentrierten Veranstaltungsreihe vorlag, unabhängig von der studienbedingten Maßnahme. Zudem muss festgestellt werden, dass derzeit keine substantiellen, aktivierenden Maßnahmen zur Unterstützung der Mitarbeit der Studierenden in den Vorlesungen enthalten waren. Somit bleibt die Herausforderung der Aktivierung der Studierenden, um sie letztlich als aktive Partner zu gewinnen, weiterhin bestehen. In Zusammenschau der aktuellen Ergebnisse und der Literatur kann somit die Hypothese aufgestellt werden, dass die Studierenden möglicherweise auf Grund der weiterhin eher passiven Rolle der Lernenden und des lehrendenzentrierten Designs der Vorlesungseinheiten trotz Überarbeitung der Vorlesungsunterlagen keinen signifikant höheren Lernzuwachs hatten. Um dies näher beurteilen zu können, sollten weitere Studien durchgeführt werden. Hierfür sollten ähnlich zum aktuellen Studiendesign alle Vorlesungen in der neuen Struktur ausgearbeitet werden, und zudem bei der Hälfte der Vorlesungen aktivierende Maßnahmen eingebunden werden. Im Anschluss sollten die Evaluationsergebnisse hinsichtlich der neuen studienbedingten Maßnahme, also der Integration von aktivierenden Maßnahmen, ausgewertet werden.

Die aktuelle Studie hat mehrere Limitationen. Die absolute und prozentuale Teilnahme an den Evaluationen lag zwar auf einem akzeptablen Niveau, wobei dennoch bemerkt werden muss, dass eine größere Teilnahme noch aussagekräftigere Ergebnisse geliefert hätte. Zudem wurden multiple statistische Tests durchgeführt (ein Test pro Frage), was zu einer alpha-Fehler-Kumulierung führen kann. Des Weiteren ermöglicht die Evaluation der Veranstaltungen mittels 5-Punkte Likert Skalen nur eine begrenzte Möglichkeit der statistischen Auswertung, da die Skala mit 5 Auswahlmöglichkeiten nur eine „grobe“ Auflösung mit schlechter Trennschärfe zwischen den zwei untersuchten Untergruppen erlaubt. Um diesem inhärenten Nachteil Rechnung zu tragen, wurde erstmals die Erhebung des Net-Promotor-Scores mittels einer 10-Punkte

Likert Skala in die Evaluation integriert. Allerdings, zeigte sich auch hierfür kein signifikanter Unterschied zwischen beiden Untergruppen, weshalb fraglich ist, inwieweit eine Erweiterung der verwendeten 5-Punkte Likert Skala auf eine 10-Punkte Likert Skala für die Fragen 1.2) bis 2.9) eine Veränderung der Ergebnisse zu Folge gehabt hätte. Die grundsätzliche Verwendung des Net-Promotor-Scores, der üblicherweise in der Konsumforschung verwendet wird, zur Evaluation von akademischen Lehrveranstaltungen stellt im Rahmen dieser Studie einen experimentellen Aspekt dar. Derzeit sind uns keine weiteren Studien bekannt, die den Net-Promotor-Scores zur Evaluation von akademischen Lehrveranstaltungen verwendet haben, weshalb eine Interpretation der aktuellen Daten im Kontext der Literatur deutlich erschwert wird. In diesem Zusammenhang scheint es jedoch sinnvoll zu sein die Frage aufzuwerfen, warum der Net-Promotor-Score nicht zur Evaluation von akademischen Lehrveranstaltungen eingesetzt werden sollte. Dies gilt insbesondere für Veranstaltungen, die freiwillig besucht werden können und nicht verpflichtend im Rahmen eines festgelegten Kurrikulums besucht werden müssen. In einem solchen Szenario, haben die Studierenden die Freiheit sich aktiv für oder gegen die Teilnahme an einer Veranstaltung zu entscheiden, wobei der Net-Promotor-Score als Maß für die Weiterempfehlung der Veranstaltung eingesetzt werden kann. Aus Sicht des Autors dieser Studie, bietet der Net-Promotor-Score die Möglichkeit einer unaufwändigen Ergänzung von Evaluationsbögen, wobei jedoch die genannten Limitationen dieses experimentellen Ansatzes berücksichtigt werden müssen. Des Weiteren muss genannt werden, dass die Evaluation nicht verpflichtend war und auf Freiwilligkeit basierte. Infolgedessen ist es möglich, dass bestimmte Studierende, die z. B. ein besonders negatives oder auch positives Bild von der Veranstaltung hatten, nicht teilgenommen haben. Auf Grund der Anzahl an verfügbaren Datensätzen (n=304), darf dieser Verzerrungseffekt jedoch als eher gering angenommen werden. Zuletzt muss die Limitation der Selbsteinschätzung von den befragten Studierenden genannt werden. Bei einer Mehrzahl der Evaluationsfragen zeigte sich, dass kein signifikanter und relevanter Unterschied zwischen den überarbeiteten und nicht überarbeiteten Vorlesungen festgestellt werden konnte. Diese Auswertung spiegelt jedoch „nur“ die Selbsteinschätzung der Studierenden wider und nicht den tatsächlichen, objektiven Lernzuwachs.

Zusammenfassung

Die grundlegende Überarbeitung der bislang vorliegenden Vorlesungsunterlagen, mittels Formulierung von Lernzielen, Gliederung in einzelne thematische Lernschritte und Wiederholungseinheiten am Ende war sicher dringend notwendig und sinnvoll, um einerseits den Studierenden die Orientierung in den initial unübersichtlichen Unterlagen zu erleichtern und sich zudem auf die zu vermittelnden Lernziele zu konzentrieren. Dennoch zeigte sich kein signifikanter Unterschied der Evaluationsergebnisse zwischen Vorlesungsveranstaltungen mit strukturiert überarbeiteten Vorlesungsunterlagen und Vorlesungsveranstaltungen ohne Überarbeitung. Folglich kann die Hypothese aufgestellt werden, dass die Studierenden möglicherweise auf Grund einer

weiterhin eher passiven Rolle der Lernenden und des lehrendenzentrierten Designs und fehlender aktivierender Maßnahmen keinen signifikant höheren Lernzuwachs hatten. Daher sollte die Vermittlung von Kompetenzen und nicht nur von Inhalten in weiteren Studien untersucht werden.

Bibliographie

- DE SHAZER 1994. *Words Were Originally Magic*. 1. Auflage, New York: W. W. Norton & Company.
- HINKIN, T. R., TRACEY, J. B., ENZ, C. A. 1997. "Scale Construction: Developing Reliable and Valid Measurement Instruments", in: *Journal of Hospitality & Tourism Research*, 21:1, S. 100–120.
- MARTON, F., SÄLJÖ, R. 1976a. "On qualitative differences in learning – I: Outcome and process", in: *British Journal of Educational Psychology*, 46:1, S. 4–11.
- MARTON, F., & SÄLJÖ, R. 1976b. "On qualitative differences in learning – II: Outcome as a function of the learner's conception of the task", in: *British Journal of Educational Psychology*, 46:2, S. 115–127.
- REICHHELD, F. F. 2003. "The one number you need to grow", in: *Harvard Business Review*, 81:12, S. 46–55.
- SAMUELS, P. 2015. "Statistical Methods – Scale reliability analysis with small samples", Birmingham City University, Centre for Academic Success. (https://www.researchgate.net/profile/Peter-Samuels/publication/280936182_Advice_on_Reliability_Analysis_with_Small_Samples/links/56cf2acd08aeb52500c9854e/Advice-on-Reliability-Analysis-with-Small-Samples.pdf; Zugriff: 11.03.2021).

Moritz Innmann ist Facharzt an der Klinik für Orthopädie des Universitätsklinikums Heidelberg. Von Oktober 2018 bis Juli 2019 forschte Moritz Innmann am Ottawa Hospital Research Institute bei Prof. Dr. Paul E. Beaulé und Prof. Dr. George Grammatopoulos. Forschungsschwerpunkte internationaler Kooperationsprojekte sind die klinische Endoprothetik und insbesondere die Interaktion von degenerativen Veränderungen zwischen Wirbelsäule und Hüftgelenken. In der studentischen Lehre ist er seit 2013 im Rahmen des Heidelberger Curriculum Medicinale der Medizinischen Fakultät Universität Heidelberg engagiert. Seit 2015 ist er Dozent für die Vorlesungen „Sportmedizin I“ (Sportorthopädie) am Institut für Sport und Sportwissenschaft der Universität Heidelberg.

PD Dr. Moritz Innmann
moritz.innmann@med.uni-heidelberg.de