

Audiovisuelle Bildungsmedienformate und ihre Lernerfolge – eine Paneldatenstudie zum Nachmachen

Sönke Häselser und Stephan Schmucker

Visuelle Bildungsmedien erfahren wachsende Verbreitung in der universitären Lehre, und die Anzahl der verwendeten Medienformate steigt. Wir untersuchen empirisch den Lernerfolg dreier Formate - Lehrvideo plus Folien/Folien plus Audio/Audio - und stellen fest, dass der Lernerfolg die aufwändige Medienproduktion nicht in jedem Fall rechtfertigt. Das Untersuchungsdesign lädt zum Nachmachen und somit zum Ermitteln des geeignetsten Medienformats in einem spezifischen Lehrkontext ein.

1. Einleitung

In der universitären Lehre (und anderorts) wird die klassische Vorlesung zunehmend durch verschiedene alternative Bildungsmedien, z.B. Lehrvideos, ergänzt oder sogar ersetzt. Solche Medien sind oftmals einer breiteren Öffentlichkeit zugänglich (Verbreitung über das Internet, z.B. über YouTube), was nicht zuletzt politische Forderungen nach ihrem verstärkten Einsatz begründet (s. z.B. Hamburger Bürgerschaft 2015). Zudem sind solche Medienformate, da sie wiederholt, an einem beliebigen Ort und zu beliebiger Zeit konsumiert werden können, nutzerfreundlicher, neben einer Reihe anderer Vorteile (Karnad 2013).

Ein Lehrkonzept, das den Einsatz nicht-traditioneller Bildungsmedien absolut notwendig macht, ist das Modell des sogenannten Inverted Classroom. Darin wird die Lehrsituation insofern umgedreht, als dass die eigentlichen Inhalte per Selbststudium vermittelt werden, oftmals unterstützt durch Lehrvideos oder andere Formen visueller Bildungsmedien. Die Vorlesungszeit wird hingegen für die Vertiefung des Inhalts genutzt, zur Erörterung von Fragen zum Inhalt, für Übungsaufgaben und generell für eine verstärkte Interaktion zwischen Lehrperson und Studierenden, sowie unter den Studierenden.

Der Einsatz visueller Bildungsmedien verbreitet sich kontinuierlich (Fischer & Spannagel 2012) und nimmt immer vielfältigere Formen an: von einfachen vertonten Vortragsfolien über Videos im Stile der Khan Academy und Vorlesungsmitschnitte bis hin zu aufwendigen, animierten ‚picture-in-picture‘-

Videos. Allerdings mangelt es gemäß Chen und Wu (2015) und vielen anderen an solider empirischer Evidenz zur Frage, welches Lehrmedienformat in einer bestimmten Bildungssituation (Unterrichtsfach, Art der Bildungseinrichtung, usw.) am geeignetsten ist, wobei die Eignung typischerweise am Lehrerfolg und/oder der Akzeptanz durch die Rezipienten gemessen wird. Chen und Wu (2015) geben einen kurzen Überblick über einige empirische Studien zu den Leistungseffekten verschiedener visueller Lehrmedien, ohne zu einem klaren Urteil zu kommen. Fey (2002) untersucht explizit, ob es einen Unterschied im Lehrerfolg von audiovisuellen Medien gegenüber reinen Audio-Medien gibt - ebenfalls ohne signifikantes Ergebnis. Interessanterweise zeigt jedoch ihre Umfrage unter den Rezipienten, dass diese audiovisuelle Medien für effektiver halten.

Eine Lehrperson, die ihren Unterricht zusätzlich mit visuellen Bildungsmedien anreichern möchte, kann also aus der Literatur wenig Orientierungshilfe bezüglich des optimalen Medienformats beziehen. Auch wird sie, wenn sie experimentell selbst die beste Lösung für die vorliegende Anwendungssituation ermitteln möchte, in Bezug auf die Untersuchungsmethoden kaum auf bestehende Studien zurückgreifen können. Denn solche Studien bauen zumeist auf Forschungsdesigns, die sich außerhalb der Forschung nicht replizieren lassen: Umfragen, Laborexperimente, spezielle Lehr- oder Testinstrumente bis hin zur Elektroenzephalografie (Chen & Wu 2015) und ähnlich aufwendige Methoden.

Die vorliegende Untersuchung möchte dieser hypothetischen Lehrperson auf zwei Weisen helfen. Erstens liefern wir zusätzliche Evidenz bezüglich der unterschiedlichen Lerneffekte und der Akzeptanz der Rezipienten von drei verschiedenen Medienformaten: Lehrvideo plus Folien; Folien plus Audiospur des Videos; nur Audio. Die Ergebnisse, die auf Paneldaten aus einem Experiment mit Studierenden an der Universität Hamburg und auf einer ‚Fixed Effects‘-Schätzung beruhen, unterstützen Feys (2002) Warnung, dass der erhebliche Produktionsaufwand audiovisueller Medien den Lernerfolg möglicherweise nicht wert ist.

Falls zweitens die Lehrperson die spärlichen aus der Literatur zu beziehenden Hinweise nicht auf ihre konkrete Lehrsituation übertragen kann oder will, so zeigt ihr diese Arbeit ein Untersuchungsdesign auf, das sich praktisch überall anwenden lässt - bei vertretbarem Mehraufwand. Indem sie dem hier geschilderten Vorgehen folgt, kann die Person im Laufe der ersten zwei oder drei

Veranstaltungen selbst herausfinden, welches der zahlreichen zur Verfügung stehenden Medienformate in ihrer spezifischen Lehrsituation die besten Ergebnisse verspricht, um sich dann im weiteren Verlauf des Semesters auf dieses Format konzentrieren zu können.

Die einzige technische Voraussetzung für die Durchführung eines solchen Experiments ist der Zugriff auf einen Klassensatz sogenannter ‚Clicker‘, der Übertragungsgeräte eines Hörsaal-Abstimmungssystems (s. z.B. Kundisch et al. 2013). Solche Systeme erfreuen sich wachsender Beliebtheit; ihnen wurden zahlreiche Vorteile in der Lehre bescheinigt (Kay & LeSage 2009, Caldwell 2007, Schmucker 2015, Schmucker & Häselser 2015, Simpson & Oliver 2007). Für unseren Zweck relevant ist die Möglichkeit, mit Clickern Single- oder Multiplechoice-Leistungstests zum Lehrinhalt durchzuführen. Dabei werden die Fragen und die dazugehörigen Antwortmöglichkeiten für alle sichtbar projiziert, die Studierenden wählen die Antworten aus und übertragen ihre Wahl mittels des Clickers an einen Computer, der sie zur späteren Auswertung speichert. Sofern genügend Clicker zur Verfügung stehen, ist diese Art von Leistungstest tatsächlich leichter durchzuführen als Varianten, bei denen die Studierenden ihre Antworten entweder auf Papier oder direkt in einen Computer eingeben. Den Auswertungsmöglichkeiten sind anschließend keine Grenzen gesetzt; wir stellen hier lediglich einige der Optionen vor.

Der Artikel gliedert sich wie folgt: Abschnitt 2 widmet sich dem Forschungsdesign. Anschließend werden die Ergebnisse vorgestellt, und zwar in Bezug auf deskriptive Statistik, Inferenzen über die Leistungseffekte der drei Medienformate und die diesbezüglichen Einstellungen der Studierenden. Abschnitt 4 reflektiert kritisch die Ergebnisse und die Datengrundlage; Abschnitt 5 fasst die Studie abschließend zusammen.

2. Forschungsdesign

Die größte Herausforderung in der Bewertung des Einflusses verschiedener Typen von Bildungsmedien auf den Lernerfolg liegt darin, weitere mögliche Einflussfaktoren auf die Leistung der Studierenden im Test auszuschließen oder konstant zu halten, die in einer Regressionsanalyse (praktischerweise) nicht oder nicht genügend berücksichtigt werden können. Zu solchen Faktoren zählen die generelle Leistungsbereitschaft der einzelnen Studierenden, ihr vorheriges Fachwissen, ihr Verhalten in der speziellen Testsituation, ihre Teil-

nahme an den Vorlesungen und am Leistungstest selbst (Selektionseffekt) usw.

Das Experiment wurde durchgeführt in drei aufeinanderfolgenden Lehrveranstaltungen im Kurs „Einführung Personal“ des Bachelorstudiengangs „Sozialökonomie“ an der Universität Hamburg. Die Studierenden wurden gebeten, jeweils in der Woche vor einem „Testtag“ genau einen von drei Typen von Lehrmedien zu konsumieren:

- Typ A ist je ein Video für jeden Testtag von etwa 8 bis 10 Minuten Länge. Das Bild ist geteilt zwischen einer Videoaufnahme des Dozenten, der den Lehrinhalt in einer improvisierten Studioumgebung erklärt, und den Lehrfolien, deren Inhalt sich mit dem mündlichen Vortrag aufbaut.
- Typ B entspricht Typ A, jedoch entfällt das Videobild des Dozenten. Die Studierenden hören also nur die Tonspur des Videos und sehen die Folien, nicht aber die Person des Dozenten.
- Typ C hat keine visuelle Komponente, sondern besteht nur aus einer Audio-Datei mit der Stimme des Dozenten aus dem Lehrvideo.

In der Lehrveranstaltung, die dem ersten Testtag vorausging, wurden den Studierenden der Zweck und Ablauf des Experiments erklärt, und sie wurden gebeten, zu einer beliebigen Zeit vor oder am ersten Testtag genau einen bestimmten der drei Typen von Bildungsmedien zu konsumieren. Die Studierenden wurden den drei Typen in einer annähernd zufälligen Weise zugeordnet, nämlich über die letzten Stellen ihrer Matrikelnummer. Dadurch wurde verhindert, dass sich die Studierenden ihren Medientyp selbst aussuchen, was die Ergebnisse invalidiert hätte. Direkt im Anschluss an diese Einführungsveranstaltung wurden die drei Mediendateien für den ersten Testtag auf der „Lecture2Go“-Plattform der Universität Hamburg unter einem den Studierenden bekannten Link veröffentlicht. Am Ende des ersten und zweiten Testtages wurden die Studierenden jeweils wieder einem Medientyp zugeordnet, mit dem Ziel, dass Studierende, die an allen drei Testtagen teilnehmen, jeden Typ genau einmal konsumieren.

Somit erhalten wir einen Paneldatensatz mit einer zeitlichen (Testtage) und einer Querschnittsdimension (Studierende). Dadurch können wir die Leistungseffekte der Medientypen mit Fixed Effects schätzen: Studierende, die an mindestens zwei Testtagen teilgenommen und mindestens zwei verschiedene

Medientypen konsumiert haben, werden jeweils einer eigenen Dummy-Variablen zugeordnet. Das sorgt dafür, dass in der Schätzung alle möglichen studierendenbezogenen Einflussfaktoren auf die Leistung konstant gehalten werden und etwaige Leistungsunterschiede also nur auf die verschiedenen Medientypen zurückzuführen sind - und möglicherweise auf Unterschiede in der Schwierigkeit der Fragen an den drei Testtagen, deshalb werden zusätzlich Dummy-Variablen für den zweiten und dritten Testtag eingeführt. Ohne diese Fixed Effects-Strategie wäre es notwendig, sehr umfangreiche Daten über die persönlichen Merkmale der Studierenden zu erheben – was dem Untersuchungsziel einer überall replizierbaren Erhebung entgegenstünde.

An den Testtagen wurden Clicker an die Studierenden ausgegeben, die damit zunächst die letzten fünf (von insgesamt sieben) Stellen ihrer Matrikelnummer übermitteln sollten. Fünf Stellen genügen in diesem Fall theoretisch, um die in diesem Kurs eingeschriebenen Studierenden eindeutig über die drei Testtage hinweg zu identifizieren, was für die Zuordnung einer Dummy-Variablen zwingend notwendig war. Anschließend wurden die Studierenden gebeten, über ihren Clicker die folgenden Angaben zu übermitteln:

1. Welchen Medientyp bzw. welche Medientypen sie im Laufe der vorangegangenen Woche tatsächlich konsumiert hatten - möglicherweise in Abweichung von dem ihnen zugeordneten Typ. Auswahlmöglichkeiten: Typen A, B und C; alle möglichen Kombinationen der drei Typen; "kein Konsum".
2. Wie viele Tage seit dem letzten Konsum eines der Bildungsmedien vergangen waren. Auswahlmöglichkeiten: 0 (Konsum am Testtag) bis 7 Tage.
3. Wie oft sie die Lehrmedien konsumiert hatten. Auswahlmöglichkeiten: ein- bis viermal oder „öfter“.

Abschließend sollten die Studierenden an jedem Testtag fünf fachliche Single-Choice-Fragen zum Inhalt der Bildungsmedien mit je drei bis fünf Antwortmöglichkeiten beantworten. Die Durchführung des gesamten Ablaufs beanspruchte etwa 30 Minuten an jedem Testtag.

3. Ergebnisse

3.1. Deskriptive Statistik

Insgesamt wurden an den drei Testtagen 174 Tests (mitunter nur teilweise) absolviert (die „gesamte Stichprobe“). Durch die Übertragung der Matrikelnummer war es möglich, 40 Studierende zu identifizieren, die an mindestens zwei Testtagen teilgenommen hatten. Dies sind die Studierenden, deren Testergebnisse für eine Fixed-Effects-Schätzung nutzbar sind. Diese Gruppe absolvierte insgesamt 96 Tests (die „nutzbare Stichprobe“).

Abbildung 1 zeigt die Häufigkeitsverteilung der in der nutzbaren Stichprobe konsumierten Medientypen. Erfreulicherweise sind die drei Typen etwa gleich oft konsumiert worden – die Studierenden haben sich also weitgehend an den ihnen zugeordneten Typ gehalten, so dass signifikante Selbstselektion ausgeschlossen werden kann. In nur zwei Fällen wurde mehr als ein Medienformat konsumiert, hieraus lassen sich also keine Schlüsse ableiten. Es erweist sich hingegen als Glücksfall, dass entgegen den Instruktionen in zahlreichen Fällen keine Lehrmedien vor dem jeweiligen Test konsumiert wurden. Diese Gruppe von Ergebnissen ist wertvoll als Benchmark zur Beantwortung der Frage, ob überhaupt eines der drei Formate einen Lerneffekt erzielen konnte. Ohne diese Gruppe wären nur Aussagen über die unterschiedlichen Effekte der drei Typen untereinander möglich gewesen.

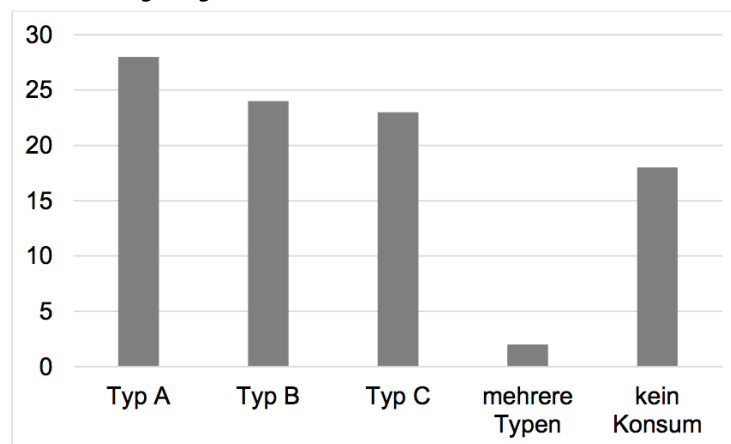


Abbildung 1: Häufigkeiten der konsumierten Medienformate

Wie erwähnt, sollten sich die Studierenden auch über den Zeitpunkt des Medienkonsums äußern. Die Frage danach verfolgte zwei Ziele: Erstens sind die Medienkonsumgewohnheiten der Studierenden an sich von Interesse. Zweitens sollte untersucht werden, ob der Zeitpunkt des Konsums mit seiner Lernwirkung korreliert - plausibel wäre, dass zeitlich naher Konsum im Test besser erinnert wird. Abbildung 2 zeigt die Häufigkeitsverteilung der Antworten. Der Konsum fand selten mehr als einen Tag vor dem Testtag statt, deshalb sind diese Ereignisse in der Kategorie „früher“ zusammengefasst. Die meisten Studierenden schauten oder hörten sich die Mediendateien am Vortag des Tests an. Bezüglich des zweiten Ziels, des Lerneffekts des Konsumzeitpunkts, war es leider nicht möglich, diese Variable in der Regressionsanalyse zu berücksichtigen, da die zahlreichen unvollständigen Antworten der Studierenden zu dieser Frage die nutzbare Stichprobe in einer Fixed-Effects-Schätzung übermäßig verkürzt hätten. Dennoch lässt eine Random-Effects-Schätzung über die gesamte Stichprobe ohne Dummy-Variablen für die Studierenden erahnen, dass zeitlich naher Medienkonsum tatsächlich gute Testergebnisse begünstigt. (Auf die umfassende Darstellung der Schätzungsergebnisse wird verzichtet.)

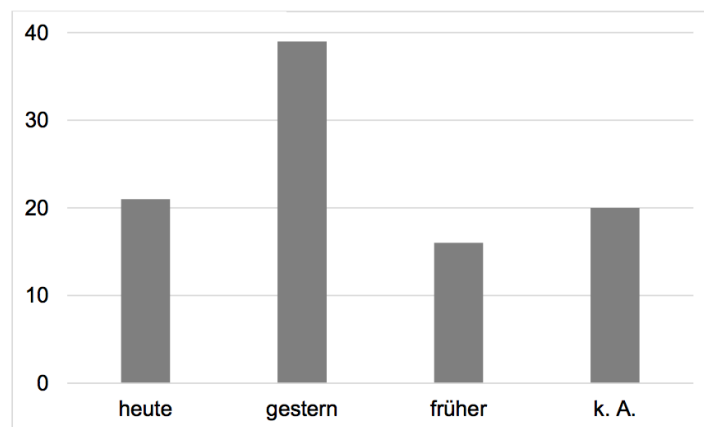


Abbildung 2: Zeitpunkt des Medienkonsums, bezogen auf den Testtag

Die letzte Frage an die Studierenden vor dem eigentlichen Wissenstest betraf die Häufigkeit des Medienkonsums. Die Ergebnisse wurden in eine Dummy-Variablen übertragen (wiederholter Konsum – ja/nein), da mehr als zweimaliger

Konsum sehr selten angegeben wurde. Nach dieser Umformulierung entstanden 65 Testergebnisse nach einmaligem Medienkonsum und 15 nach mehrfachem Konsum; in den verbleibenden 16 Fällen wurde keine Antwort gegeben. Bedauerlicherweise eignete sich auch diese Variable – aus den gleichen Gründen wie oben – nicht zur Schätzung in einer Fixed-Effects-Regression. In einer Random-Effects-Schätzung über die gesamte Stichprobe zeigt sich jedoch, dass wiederholter Medienkonsum erwartungsgemäß zu signifikant besseren Lernergebnissen führt (die detaillierten Ergebnisse sind wiederum nicht aufgeführt).

3.2. Leistungseffekte

Die wichtigsten Ergebnisse einer Regression zum Einfluss der verschiedenen Typen von Bildungsmedien auf den Lernerfolg sind in Tabelle 1 zusammengefasst. Die Anzahl der an einem Testtag korrekt beantworteten Fragen als abhängige Variable wird hier in Verbindung gebracht mit je einer Dummy-Variablen für die drei Medientypen („kein Konsum“ als Referenzkategorie), mit Dummy-Variablen für den zweiten und dritten Testtag (Testtag 1 als Referenzkategorie) und mit einem Satz von Dummy-Variablen für die 40 Studierenden in der nutzbaren Stichprobe. In dieser Fixed-Effects-Schätzung wurde Poisson-Regression angewandt, um dem Umstand Rechnung zu tragen, dass die abhängige Variable nur diskrete Werte annimmt. Die geschätzten Koeffizienten für die einzelnen Studierenden-Dummies sind in der Tabelle nicht wiedergegeben, einige sind allerdings hochsignifikant. Die Variablen zu Medienkonsumzeitpunkt und -häufigkeit konnten aus den oben genannten Gründen in diesem Modell nicht berücksichtigt werden.

Tabelle 1: Robuste Poisson Regression mit Fixed Effects für die Studierenden. 96 Beobachtungen von 40 Studierenden.

Variable	Regressionskoeffizient	P > z
Medientyp A	0,5339	0,020
Medientyp B	0,6338	0,007
Medientyp C	0,3726	0,163
mehrere Typen	1,1984	0,003
Testtag 2	-0,3808	0,002
Testtag 3	-0,0145	0,891

Die Ergebnisse zeigen, dass die Studierenden relativ zum ersten Testtag am zweiten Tag signifikant schlechter abschnitten, aber nur marginal schlechter am dritten. Dies spiegelt wahrscheinlich unabsichtliche Differenzen in der Schwierigkeit der Fragen an den drei Tagen wider. Das Hauptinteresse gilt jedoch den Koeffizienten der drei Medientypen. Der Konsum von allen drei Formaten scheint das Abschneiden im Test zu begünstigen; bei Typen A und B ist dieser Effekt deutlich signifikant, bei Typ C kann die Nullhypothese bezüglich des Leistungseffekts hingegen nicht zurückgewiesen werden. Der Schätzwert zu „mehreren“ Typen verdient aus zwei Gründen keine weitere Interpretation: Erstens liegen diesem Wert wie erwähnt nur zwei Testergebnisse zugrunde. Zweitens ging diesen Ergebnissen logischerweise mehr als einmaliger Medienkonsum voran, und die Konsumhäufigkeit ist ansonsten im Modell nicht berücksichtigt. Wir könnten also nicht unterscheiden, ob der augenscheinliche Leistungseffekt auf die erhöhte Konsumhäufigkeit zurückzuführen ist oder auf die Tatsache, dass unterschiedliche Medientypen konsumiert wurden. Als abschließende Beobachtung zu Tabelle 1 halten wir das Ranking der Koeffizienten zu den drei Medientypen fest: B>A>C.

Unterscheiden sich die Lerneffekte der drei Medientypen signifikant voneinander? Dazu führen wir paarweise χ^2 -Tests zur Gleichheit der entsprechenden Regressionskoeffizienten durch. Die Ergebnisse werden in Form der jeweiligen Wahrscheinlichkeitswerte in Tabelle 2 dargestellt.²⁰ Allein der Unterschied

²⁰ Die Werte stehen für die Wahrscheinlichkeit, allein durch zufällige Stichprobenzie-

zwischen den Lerneffekten der Typen B und C scheint signifikant zu sein, allerdings auch nur, wenn wir ein Signifikanzniveau von 10 Prozent anlegen. Deutlichere Ergebnisse würden mutmaßlich eine bessere Datenbasis erfordern.

Tabelle 2: Paarweise X²-Tests der entsprechenden Regressionskoeffizienten der drei Medientypen

Medientypen	B	C
A	0,295	0,257
B	-	0,054

4. Einschätzung der Studierenden

Am Ende des dritten Testtages wurden die Studierenden gebeten, zu zwei Zusatzfragen bezüglich ihrer Einstellungen zur mediengestützten Lehre Stellung zu nehmen. Die Erste davon fragte nach dem Grad der Zustimmung zu der Aussage „Ich empfinde die Vorbereitung mit audiovisuellen Medien als sinnvoll.“ Abbildung 3 zeigt die Häufigkeitsverteilung der Antworten.²¹ Offensichtlich begrüßen fast alle Studierenden den Einsatz von audiovisuellen Bildungsmedien. Dieses Ergebnis überrascht nicht, wenn man sich die spezifischen Umstände des Medieneinsatzes in diesem Anwendungsbeispiel vor Augen führt: Die Mediendateien wurden den Studierenden nicht als Alternative, sondern eher als Ergänzung zu den weiterhin stattfindenden regulären Vorlesungen angeboten. Nach dem Abschluss jedes Tests wurden dessen Inhalte in der Vorlesung diskutiert, so dass die Studierenden die relevanten Begriffe und Zusammenhänge auf diese Weise auch dann erlernen konnten, wenn sie die Lehrmedien nicht konsumiert hatten. Der Konsum, ebenso wie der Nicht-Konsum, war für sie kaum mit Kosten verbunden. Die Medien stellten für sie eine zusätzliche, optionale Lernquelle dar - die sie deshalb vorhersehbarerweise begrüßten. Die Verteilung der Antworten zu dieser Frage sähe wohl anders aus, wären die Studierenden beispielsweise vor die Entscheidung

hung aus einer Population, in der die Koeffizienten gleich sind, Differenzen zwischen den Koeffizienten zu erhalten, die größer sind als die tatsächlich aus der Regression resultierenden Differenzen.

²¹ Abbildungen 3 und 4 umfassen die Antworten von allen teilnehmenden Studierenden, unabhängig davon, ob ihre Testergebnisse in der Regression berücksichtigt werden konnten.

zwischen Vorlesung und Bildungsmedien als strikte Alternativen gestellt worden.²² Zudem ist zu bedenken, dass die Einstellungen der Studierenden nur bedingt als Grundlage für die Entscheidung über den Medieneinsatz dienen können. Schließlich haben die Studierenden die erheblichen Kosten des Einsatzes nicht zu tragen. Aus ihrer Zustimmung kann also nicht direkt eine Entscheidung für einen Einsatz abgeleitet werden; ihre Ablehnung hingegen könnte ein gewichtiges Argument in einer Entscheidung gegen einen Einsatz sein.

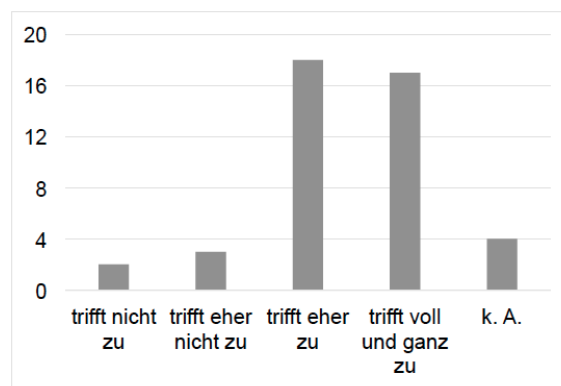


Abbildung 3: „Ich empfinde die Vorbereitung mit audio-visuellen Medien als sinnvoll.“

Die zweite Zusatzfrage verspricht noch interessantere Einsichten. Die Frage an die Studierenden lautete „Welcher Lehrvideo-Typ hat Ihren Lernfortschritt am besten unterstützt?“ Abbildung 4 zeigt die Ergebnisse. Die grauen Säulen entsprechen den ungewichteten Häufigkeiten. Sie zeigen eine deutliche Präferenz für Typ B, der mehr als doppelt so viele „Stimmen“ erhält wie Typ A. Nur zwei Studierende bevorzugten Typ C, die reinen Audiodateien. Dieser erste Eindruck wird noch deutlicher, wenn wir die Anzahl der Antworten gewichten mit der Anzahl der verschiedenen Medientypen, die die Antwortenden insgesamt im Laufe des Experiments konsumiert haben (schwarze Säulen). Hintergrund ist, dass die Aussage von jemandem, der drei verschiedene Formate

²² Fischer und Spannagel (2012) zählen eine Reihe von Möglichkeiten auf, wie audiovisuelle Medien im Verhältnis zu herkömmlichen Vorlesungen eingesetzt werden können. In einem dem unsrigen nicht unähnlichen Versuchsaufbau stellen auch sie einen starken Zuspruch der Studierenden für elektronische Lehrmedien fest.

erfahren hat, wohl gewichtiger ist als diejenige von jemandem, der tatsächlich kaum einen validen Vergleich anstellen kann, da er nur einen Typ konsumiert hat - oder gar keinen. In jedem Fall stellen wir fest, dass das Ranking der Präferenzen der Studierenden übereinstimmt mit dem Ranking, das sich aus den Leistungstests ergibt.

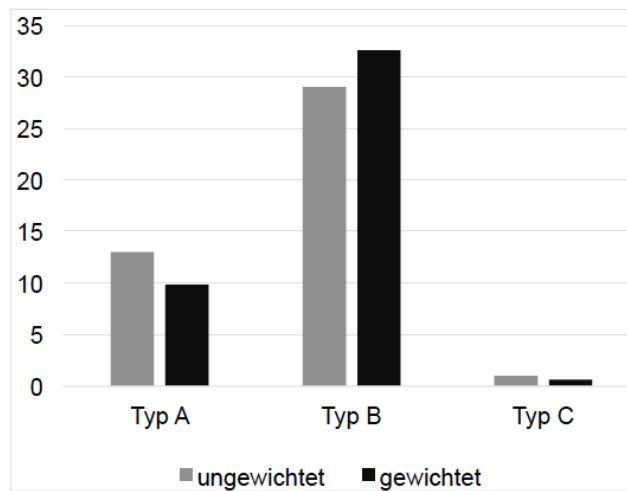


Abbildung 4: „Welcher Lehrvideo-Typ hat Ihren Lernfortschritt am besten unterstützt?“

5. Kritische Reflexion

Wie gezeigt, legen die Ergebnisse der Leistungstest und die der Umfrage nahe, dass die visuellen Bildungsmedien des Typs B (Audio + Folien) bessere Lernerfolge erzielen als solche des Typs A (zusätzliche Videoaufzeichnung) und dass diese Wirkung von den Studierenden auch entsprechend eingeschätzt wird. Diese Kongruenz ist ermutigend, allerdings ist das Ergebnis an sich zunächst überraschend und steht im Kontrast zu entgegengesetzten Resultaten einer Studie von Fey (2002) in einem vergleichbaren Kontext. Zunächst wäre doch anzunehmen, dass der zusätzliche Informationsgehalt des Formats A gegenüber B von den Studierenden gewürdigt werden und sich auch in den Testergebnissen niederschlagen sollte. Sind nicht mehr Informationen immer besser als weniger, gilt die in der Mikroökonomik übliche Annahme des „free disposal“ hier (oder gar generell in Bezug auf Informationen) nicht? Anscheinend nicht.

Die Ergebnisse mögen Ausdruck des Umstands sein, dass multimodale Bildungsmedien das Risiko eines „cognitive overload“ steigern, also die Gefahr der Unfähigkeit, alle dargebotenen Informationen zu verarbeiten. Zusätzlicher Input kann dann die Lernwirkung sogar behindern (Sweller et al. 1998). Videobilder fördern das Lernen indem - und nur insoweit als dass - sie die Lernenden motivieren und Emotionen in ihnen hervorrufen (Edelmann 2000). Ob die hier verwendeten Videos des Typ A geeignet waren, diese Ziele zu erreichen, bleibt offen.

Neben solchen Interpretationen des Inhalts sind an dieser Stelle auch einige Überlegungen zu der Aussagekraft der Ergebnisse angebracht, welche in erster Linie von der Qualität der zugrundeliegenden Daten abhängt. Die aus der Software, die die Clicker-Antworten aufzeichnet, exportierten Daten zeigen leider einige Anzeichen von eingeschränkter Verlässlichkeit. In einer Reihe von Fällen gaben die Studierenden offensichtlich unvollständige oder falsche Antworten ab, sei es aus Absicht (z.B. weil sie über die Matrikelnummer ihre Identität nicht preisgeben wollten) oder aus Versehen (z.B. weil sie unbeabsichtigt eine falsche Clicker-Taste oder die richtige Taste nicht fest genug gedrückt haben). Diese Mutmaßungen über das Verhalten der Studierenden stützen sich u.a. auf folgende Beobachtungen:

- Einige der Studierenden gaben weniger als die erbetenen fünf Stellen ihrer Matrikelnummer an; einige gaben sogar überhaupt keine Informationen preis, die ihre Identifikation über die drei Testtage hinweg ermöglicht hätte. In einigen Fällen handelte es sich hierbei sicherlich um absichtliches Verhalten. In anderen Fällen nahmen solche Studierenden hingegen an zwei oder drei Testtagen teil und ließen eine oder mehrere der geforderten Stellen an nur einem der Testtage aus, was auf eine unabsichtliche Auslassung schließen lässt. Sofern drei oder vier Stellen übertragen wurden, war es mitunter möglich, die volle Matrikelnummer zu rekonstruieren und dadurch das dazugehörige Testergebnis solchen von anderen Testtagen zuzuordnen und so die nutzbare Stichprobe zu vergrößern. Jedoch barg dieses Vorgehen eine gewisse Fehlergefahr in Fällen, in denen Studierende an den Tests teilnahmen, ohne sich in den Kurs eingeschrieben zu haben.
- In 13 Fällen der gesamten Stichprobe gaben Studierende an, sie hätten keines der Medienformate konsumiert, obwohl sie zwei Fragen später einen Zeitpunkt des Konsums angaben, anstatt auf diese Frage

gar nicht zu antworten. Ähnliche Widersprüche bestanden bei der Frage zur Konsumhäufigkeit.

- Einige Studierende gaben in der entsprechenden Zusatzfrage eine persönliche Einschätzung zur Wirksamkeit der Medientypen ab, obwohl sie nach eigener Aussage keinen davon konsumiert hatten.

Unvollständige Angaben der Studierenden führen zu einer Verkleinerung der nutzbaren Stichprobe; falsche Angaben führen zu (hoffentlich zufälligen) Störungen in den Daten, wodurch die Standardfehler der geschätzten Regressionskoeffizienten steigen. In beiden Fällen sinkt die Wahrscheinlichkeit, statistisch signifikante Ergebnisse zu finden. Dies sollte bei der Interpretation der Ergebnisse berücksichtigt werden: Bei vollständigeren und qualitativ besseren Daten wären wohl deutlichere Effekte zu erwarten gewesen. Die Verbesserung der Datenbasis stellt, neben einer stärkeren theoretischen Fundierung, somit eine Hauptaufgabe für zukünftige Forschung in dieser Richtung dar.

In einer zweiten Durchführung dieses Experiments müssten wohl auch die Fragen und Antwortmöglichkeiten einer Revision unterzogen werden. Bei einer am ersten Testtag gestellten Frage erhielt eine der möglichen Antworten 84% der „Stimmen“, alle übrigen gingen an die zweite Antwort und also keine an die dritte. Diese Frage konnte somit nur sehr wenig zur Einschätzung der Leistung der Studierenden beitragen. Zudem führte ein kleines Missgeschick am dritten Tag dazu, dass eine weitere Frage praktisch wertlos wurde: Eine der fünf Antwortmöglichkeiten war versehentlich leicht anders formatiert dargestellt als die anderen vier, was anscheinend fast alle Studenten (korrekt) als Hinweis interpretierten, dass es sich hierbei um die richtige Antwort handelte. Diese ungewollten Erleichterungen am ersten und dritten Testtag mögen erklären, warum die Studierenden an diesen Tagen jeweils signifikant mehr richtige Antworten erzielten als am zweiten Testtag (s. Tabelle 1).

6. Zusammenfassung und Ausblick

Man versetze sich in die Lage einer Lehrperson, die ihren Kurs durch den Einsatz visueller Bildungsmedien bereichern möchte, sich aber nicht sicher ist, welches der vielen vorhandenen Formate für ihre spezifische Anwendungssituation geeignet ist. Von der vorhandenen Literatur kann sie wenig Hilfestellung erwarten: Erstens sind die Ergebnisse uneinheitlich. Zweitens ist unklar, inwieweit sich unter Laborbedingungen erzielte Resultate und die Lehren

daraus auf ihren speziellen Kontext übertragen lassen. Drittens kann sie die in der Literatur angewendeten, oftmals extrem aufwendigen Untersuchungsdesigns unmöglich replizieren, um etwa die gesuchten Antworten anhand ihrer eigenen Lerngruppe selbst empirisch zu finden.

In diesem Artikel wurde hingegen ein verhältnismäßig einfaches Untersuchungsdesign vorgestellt, mit dem die Effektivität verschiedener Medienformate in einer großen Bandbreite von Kontexten durch die Lehrperson selbst untersucht werden kann. Das Vorgehen hat sich als weitgehend robust erwiesen gegenüber nur sporadischer Teilnahme der Studierenden, gegenüber mitunter eher halbherzigem Mitwirken und anderen Hindernissen des realen Lehrbetriebs. Die Voraussetzungen der Durchführung beschränken sich auf eine hinreichende Gruppengröße (50+), einen Klassensatz Clicker und gewisse Statistikenkenntnisse.

Neben dieser praxisorientierten Anleitung leistet die vorliegende Arbeit auch einen Beitrag zur Literatur im Sinne eines Tests des didaktischen Mehrwerts visueller Bildungsmedien. Zudem ermöglicht sie eine Aussage über den Mehrwert von Multimodalität, denn die drei untersuchten Medienformate - nur Audio/Text (Folien) plus Audio/Text plus Video - repräsentieren jeweils eine unterschiedliche Anzahl von Informationsmodi (eins bis drei).

Das Experiment wurde durchgeführt unter Bachelor-Studierenden der Sozialökonomie im Kurs „Einführung Personal“ an der Universität Hamburg. Die Ergebnisse zeigen, dass der vorherige Konsum eines Lehrvideos mit Folien (Typ A) bzw. der Folien begleitet von der Stimme des Dozenten (Typ B) jeweils signifikant das Lernergebnis der Studierenden verbessern, wobei die Effekte bei Typ B noch stärker sind. Kein signifikanter Lernerfolg wird hingegen durch eine auf dem Lehrvideo basierende reine Audiodatei (Typ C) erzielt. Dieses Ergebnis spiegelt sich auch wider in der subjektiven Einschätzung der Studierenden, welches der drei Formate ihren Lernfortschritt am besten unterstütze.

Dass die eigentlichen Videos in beiden Betrachtungsweisen marginal schlechter abschneiden als die Variante, in der zu den Folien nur die Stimme des Dozenten zu hören ist, überrascht zunächst. Das Ergebnis legt nahe, dass Lehrvideos ihr Ziel der motivierenden und sogar emotionalisierenden Ansprache erreichen müssen, damit das Risiko des „cognitive overload“ durch das zusätzliche bewegte Bild aufgewogen wird. In jedem Fall stimmen wir Fey (2002) in der Warnung zu, dass der etwaige didaktische Mehrwert von Lehrvideos gegenüber z.B. „gesprochenen“ Folien den erheblichen zusätzlichen

Produktionsaufwand des Videoformats oftmals nicht rechtfertigt. Auch warnen wir davor, diese – oder irgendwelche – Ergebnisse übermäßig zu generalisieren: Die Anzahl der vorhandenen Formate von visuellen Bildungsmedien und die Anzahl der unterschiedlichen Bildungskontexte, in denen diese Medien eingesetzt werden könnten, sind einfach zu groß für pauschale Empfehlungen. Gerade deshalb sprechen wir uns, wo immer möglich, für eine empirische ad hoc-Erhebung im spezifischen Lehrkontext aus.

Sollte dieses Experiment in Zuge zukünftiger Forschung wiederholt werden, so erscheint der Verzicht auf den dominierten Medientyp C sinnvoll, um sich auf die Unterschiede der effektiveren Typen A und B konzentrieren zu können. Zudem sollten Anstrengungen in Richtung eines vollständigeren und qualitativ besseren Datensatzes unternommen werden. Eine mögliche Maßnahme zu diesem Zweck ist z.B. die Auslobung kleiner Bargeldpreise für engagierte Teilnahme.

Autoren

Dr. Sönke Häsel
Freiberuflicher Forscher
E-Mail: soenke.haeseler@gmail.com

Dr. Stephan H. Schmucker
Wissenschaftlicher Mitarbeiter
Universität Hamburg
E-Mail: Stephan.Schmucker@wiso.uni-hamburg.de

Literatur

- Caldwell, J.E. (2007). Clickers in the large classroom: Current research and best-practice tips. *Life Sciences Education*, 6(1), 9-20.
- Chen, C.-M., Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 80, 108-121.
- Edelmann, W. (2000). *Lernpsychologie*. Weinheim: BeltzPVU.
- Fey, A. (2002). Audio vs. Video: Hilft Sehen beim Lernen? Vergleich zwischen einer audiovisuellen und auditiven virtuellen Vorlesung. *Unterrichtswissenschaft*, 30(4), 331-338.
- Fischer, M., Spannagel, C. (2012). Lernen mit Vorlesungsvideos in der umgedrehten Mathematikvorlesung. In: Desel, J., Haake, J. M. and Spannagel, C. (Hrsg.), *DeLFI 2012 – Die 10. e-Learning Fachtagung Informatik*. Hagen/Heidelberg: Gesellschaft für Informatik, 225-236.
- Hamburger Bürgerschaft (2015). *Digitales Lehren und Lernen an den staatlichen Hamburger Hochschulen*. Mitteilung des Senats an die Bürgerschaft. Drucksache 20/14262
- Karnad, A. (2013). Student use of recorded lectures: a report reviewing recent research into the use of lecture capture technology in higher education, and its impact on teaching methods and attendance. London School of Economics and Political Science.
- Kay, R.H., LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53, 819-827.
- Kundisch, D. et al. (2013). Classroom Response Systems. *Informatik Spektrum*, 36(2), 389-393.
- Schmucker, S. (2015). Cognitive activation in mass lectures through electronic voting systems in the lecture theatre. In: *International Conference on Quality in Higher Education - ICQH 2014, Proceedings Book, Sakarya/Turkey*, 67-72.
- Schmucker, S., Häseler, S. (2015). Student diversity, peer instruction and classroom response systems. *The Turkish Online Journal of Educational Technology, Special Issue 2 for INTE 2015*, 508-516.
- Simpson, V., Oliver, M. (2007). Electronic voting systems for lectures then and now: A comparison of research and practice. *Australasian Journal of Educational Technology*, 23(2), 187-208.
- Sweller, J., van Merriënboer, J.J.G., Paas, F.G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296.

Online zugänglich unter:

bildungsforschung

Ausgabe 1 (2015), 12. Jahrgang

Sönke Häsel und Stephan Schmucker (2015). Audiovisuelle Bildungsmedienformate und ihre Lernerfolge – eine Paneldatenstudie zum Nachmachen. In: bildungsforschung, Jahrgang 12, Ausgabe 1, URL: <http://www.bildungsforschung.org/>