

Zweitveröffentlichung



Zmiskol, Tabea; Hess, Miriam

Multimodale Modellantworten als direktes Feedback bei der individuellen Unterrichtsvideoanalyse

Datum der Zweitveröffentlichung: 27.01.2026

Verlagsversion (Version of Record), Zeitschriftenartikel

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-112782x

Erstveröffentlichung

Zmiskol, Tabea; Hess, Miriam (2026): Multimodale Modellantworten als direktes Feedback bei der individuellen Unterrichtsvideoanalyse, in: die hochschullehre, Bielefeld: wbv Media GmbH & Co. KG, Jg. 12, Nr. 14, S. 151–159, doi: 10.3278/HSL2614W.

Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis der Rechteinhaberinnen und Rechteinhaber einholen.

Für dieses Dokument gilt eine Creative-Commons-Lizenz.



Die Lizenzinformationen sind online verfügbar:

<https://creativecommons.org/licenses/by-sa/4.0/legalcode>

die hochschullehre – Jahrgang 12 – 2026 (14)

Herausgebende des Journals: Svenja Bedenlier, Ivo van den Berk, Sarah Berndt, Jonas Leschke, Peter Salden, Antonia Scholkmann, Angelika Thielsch

Beitrag in der Rubrik Forschung

DOI: 10.3278/HSL2614W

ISSN: 2199–8825 wbv.de/die-hochschullehre



Multimodale Modellantworten als direktes Feedback bei der individuellen Unterrichtsvideoanalyse

TABEA ZMISKOL & MIRIAM HESS

Zusammenfassung

Im Diskurs um Reformen in der Lehrkräftebildung wird der Forderung nach fundierten, berufspraktisch relevanten Kompetenzen u. a. mit dem Einsatz von Unterrichtsvideos begegnet. Dabei stellt sich nicht nur die Frage nach geeigneten Maßnahmen, um den Kompetenzerwerb auch im digitalen Raum bei der individuellen Videoanalyse zu begleiten, sondern es gilt auch zu untersuchen, welche Modalität diese Begleitmaßnahmen haben sollten, um für Studierende im Sinne der *Cognitive Theory of Multimedia* ansprechend und gut nutzbar zu sein. Es wurde eine Erhebung in einem videobasierten Online-Lernsetting durchgeführt, bei dem 137 Grundschullehramtsstudierende zwischen auditiven und transkribierten Modellantworten als direktes Feedback wählen konnten. Analog zur Feedbackmodalitätsforschung bestätigte sich das Audio als präferierte Modalität. Jedoch wurde ein einschränkender Reihenfolgeeffekt festgestellt. Erste Erklärungen, warum Studierende Audio oder Transkript präferieren, liefert eine Inhaltsanalyse.

Schlüsselwörter: Unterrichtsvideoanalyse; Modellantworten; Feedbackmodalität; Audio-Feedback; Text-Feedback, Cognitive Theory of Multimedia

Multimodal model responses as direct feedback for individual teaching video analysis

Abstract

In the discourse on reforms in teacher training, the demand for well-founded, professionally relevant skills is met, among other things, with the use of teaching videos. This not only raises the question of suitable measures to accompany the acquisition of skills in the digital space during individual video analysis, but it is also necessary to investigate what modality these accompanying measures should have in order to be appealing and usable for students in the sense of the *Cognitive Theory of Multimedia*. A survey was carried out in a video-based online learning setting in which 137 primary school teacher trainees were able to choose between auditory and transcribed model responses as direct feedback. Analogous to the feedback modality research, audio was confirmed as the preferred modality. However, a limiting order effect was found. A content analysis provides initial explanations as to why students prefer audio or transcript.

Keywords: Instructional video analysis; model responses; feedback modality; audio feedback; text feedback, Cognitive Theory of Multimedia

1 Einleitung

Bei der Frage nach der geeigneten Modalität von Lernsettings öffnet sich der Blick auf eine umfangreiche Forschung. So zählt das *Modalitätsprinzip* (*Modality Principle*) zu den meist erforschten Phänomenen der *Cognitive Theory of Multimedia* (Castro-Alonso & Sweller, 2022). Es besagt, dass visuelle Instruktionen wie Diagramme und Modelle eher auditiv und nicht ebenfalls visuell via Text vermittelt werden sollten. Jedoch sind Fragen dazu offen, unter welchen Umständen das Prinzip nicht greift und wie es mit anderen Multimedia-Prinzipien zusammenhängt. In Bezug auf die Modalitäten von Feedback lässt sich festhalten, dass sich Audio-Feedback als lernwirksamer und persönlicher andeutet als Text-Feedback, doch sind die Ergebnisse hierzu nicht immer eindeutig und teilweise widersprüchlich (zsf. Alharbi, 2021). In diesem Beitrag werden Ergebnisse der Feedbackforschung und der *Cognitive Theory of Multimedia* (Mayer, 2022) auf den Einsatz von Modellantworten übertragen und erprobt.

2 Theoretisch-empirischer Hintergrund

Im aktuellen Diskurs um Reformen der Lehrkräftebildung wird seit einigen Jahren gefordert, angehenden Lehrkräften frühzeitig zu ermöglichen, theoriebasiert und situiert unterrichtsbezogene Kompetenzen zu erwerben (BMBF, 2019; Seidel, 2022). Unterrichtsvideos erweisen sich hier als besonders vielversprechend, um die sogenannte *professionelle Wahrnehmung* zu schulen (König et al., 2022), die einen positiven Einfluss auf Unterrichtsqualität und dadurch auf Leistungen von Schüler:innen ausübt (Gaudin & Chaliès, 2015; Holodynski et al., 2022). Um Studierenden eine individuelle, intensive und flexible Auseinandersetzung mit Unterrichtsvideos zu ermöglichen, bietet es sich angesichts großer Lehrveranstaltungen wie Vorlesungen an, Unterrichtsvideoanalysen als Selbstlernphasen in den digitalen Raum zu verlagern. Hierbei sollten dysfunktionale sowie funktionale Videobeispiele besprochen werden (Böhnke et al., 2022). Die Selbstlernphase kann dabei von instruktionalen Elementen (z. B. Beobachtungsfragen, Videosegmentierung) sowie Modellantworten (*worked-examples*) als direktes Feedback unterstützt werden (Huxham, 2007; Renkl, 2022). Jedoch ist noch unklar, welche Begleitmaßnahmen in welchem Kontext besonders hilfreich sind (Seidel, 2022). In Bezug auf Feedback deutet sich im digitalen Lernsetting Audio- im Vergleich zu Textfeedback als lernwirksamer an und wird von Studierenden womöglich aufgrund der wahrgenommenen *sozialen Präsenz* der Dozierenden präferiert (zsf. Alharbi, 2021). Gleichzeitig wird diskutiert, ob multimodales Feedback (z. B. Kombination von Audio- und Textfeedback), auch angesichts einer inklusiven Hochschule, zu präferieren ist (Delere & Wilkens, 2024).

3 Fragestellungen des Beitrags

Das Teilprojekt *Interaktive Unterrichtsvideos in der Lehrkräftebildung* (*InViLebi*), das im Rahmen des Projekts *Digitale Kulturen der Lehre entwickeln* (*DiKuLe*) von der Stiftung *Innovation in der Hochschullehre* gefördert wird, untersucht, wie sich sogenannte *interaktive Unterrichtsvideos* gewinnbringend in der Blended-Learning-Lehre einsetzen lassen. Ein besonderer Fokus liegt auf der Frage, wie im digitalen Raum eine sinnvolle, direkte Rückmeldung an Studierende erfolgen kann. Im bisherigen Projektverlauf wurden erfolgreich Sprachkommentare, d. h. auditiv angebotene Modellantworten, als direktes Feedback bei der Unterrichtsvideoanalyse eingesetzt (Zmiskol & Hess, 2024). Im vorliegenden Beitrag sollen nun folgende vier Fragen geprüft werden:

- Welche Modalität nutzen Studierende bei der Rezeption von Modellantworten als direktes Feedback während der Unterrichtsvideoanalyse, wenn sie die Wahl zwischen einer Audio- und einer Text-Datei haben? (F1)

- Welche Modalität präferieren Studierende nach Abschluss der Unterrichtsvideoanalyse rückblickend und hinsichtlich zukünftiger Unterrichtsvideoanalysen? (F2)
- Gibt es einen Zusammenhang zwischen der Modalitätsnutzung und der abschließenden Modalitätspräferenz mit der Art der Videoanalyse, dem Videobeispiel und mit der Länge einer Modellantwort? (F3)
- Wie begründen die Studierenden ihre abschließende Modalitätspräferenz? (F4)

4 Methodisches Vorgehen

Im Sommersemester 2024 wurde an der Universität Bamberg im Fach Grundschulpädagogik eine quasi-experimentelle Erhebung durchgeführt. Die Erhebung war als eine asynchrone Online-Selbsterlernereinheit zum Thema *Umgang mit Leistungsheterogenität* in ein Blended-Learning-Grundlagenseminar eingebunden. Innerhalb von zwei Wochen bearbeiteten insgesamt 137 Studierende die Lerneinheit (davon 86,1% weiblich¹; Fachsemester: $M = 2.6$, $SD = 1.5$, $Min = 2$, $Max = 10$).

Die Studierenden wurden randomisiert vier Interventionsgruppen zugeteilt und analysierten während der Intervention sieben Segmente ($Min = 00:39$ min, $Max = 03:25$ min) aus zwei Unterrichtsvideos, die Ausschnitte aus dem Deutschunterricht von zwei verschiedenen Lehrkräften in der ersten Klasse zeigen. Die Unterrichtsvideos entstammen der *PERLE*-Studie und zeigen somit einen weitgehend authentischen, nicht gescripteten Unterrichtsablauf (Lotz & Corvacho del Toro, 2013).

Die Studierenden der Gruppen A und C analysierten zunächst die vier Segmente des Videos, das durch ein vorausgehendes Expert:innenrating ($n = 6$) weitgehend als funktional eingestuft wurde, und im Anschluss die drei Segmente des Videos, das als eher dysfunktional eingestuft wurde. Gruppe B und D starteten dagegen mit dem dysfunktionalen Video und analysierten im Anschluss das funktionale Video. Dabei analysierten Gruppe A und B das für sie erste Video anhand von Beobachtungsfragen und das für sie zweite Videos offen, d. h. ohne Beobachtungsfragen. Die Beobachtungsfragen orientieren sich dabei an Subprozessen der *professionellen Wahrnehmung* und fordern demnach die Studierenden auf, auf bestimmte Ereignisse im Video zu achten, diese zu interpretieren und bei Bedarf Handlungsalternativen zu generieren (Blömeke, 2025; Blömeke et al., 2015). Gruppe C und D starteten dagegen mit einer offenen Videoanalyse und analysierten das für sie zweite Video fragengeleitet. Somit unterschieden sich die vier Interventionsgruppen hinsichtlich der Art der Videoanalyse (offen/fragengeleitet), der Art des Videobeispiels (funktional/dysfunktional) und deren Reihenfolge.

Zu jedem der sieben Videosegmente wurde direkt im Anschluss an die eigene Bearbeitung eine Modellantwort angeboten. Die Modellantworten wurden ebenfalls anhand des oben bereits genannten Expert:innenratings ($n = 6$) erstellt, um theoretische Bezüge und Fachbegriffe ergänzt und entsprechend der Subprozesse der professionellen Wahrnehmung strukturiert. So wird zunächst ein Ereignis im Video kurz beschrieben und anschließend interpretiert. Dabei wird erklärt, aus welchen Gründen dieses Ereignis hinsichtlich des Beobachtungsfokus *Umgang mit Leistungsheterogenität* als funktionales oder dysfunktionales Beispiel zu bewerten ist. Abschließend werden bei Bedarf Handlungsalternativen diskutiert um zu reflektieren, was auf welche Weise verbessert werden könnte. Die folgenden zwei Auszüge stammen aus zwei der vier Modellantworten (vgl. Abbildung 1: M1 und M3) zu dem Unterrichtsvideo, das als funktional eingeschätzt wurde, und zeigen exemplarisch den Aufbau der Modellantworten:

M1) Die Lehrerin [...] stellt die Schüler:innen vor die Wahl, einen „leichten“ oder einen „schweren“ Text zu lesen. Auf diese Weise realisiert sie eine innere Differenzierung. Zudem haben die Kinder durch die freie Textwahl einen Anlass, eine realistische Selbsteinschätzung auszubilden: *Welchen Text traue ich mir zu? Welcher passt am besten zu meinem Fähigkeitslevel?* Bei zunehmend realistischer Selbsteinschätzung kann der Einbezug der Schüler:innen in die Textauswahl zu einer größeren Adaptivität führen, als wenn

1 Der hohe Anteil weiblicher Studierender ist in Grundschullehrantsstudiengängen repräsentativ.

die Lehrkraft die Zuteilung ausschließlich auf Basis ihres eigenen diagnostischen Urteils beschließt, das durchaus manchmal fehlerhaft sein kann. Die Gefahr einer Unter- oder Überforderung der Kinder könnte so minimiert werden, zumindest wenn die Schüler:innen schon in der Lage sind, sich selbst gut einzuschätzen.

M3) Die Schülerin, die bereits nach kurzer Zeit wieder zur Lehrerin kommt, hatte bereits den schwierigeren Text gewählt, war damit aber offenbar dennoch noch unterfordert. Nun zieht die Lehrerin wie ein Ass aus dem Ärmel einen dritten, noch etwas schwierigeren Text als Anschlussaufgabe für das Mädchen hervor. Diesen noch schwierigeren Text scheint das Mädchen gerne als neue Herausforderung anzunehmen. Die Lehrerin lobt sie und belohnt sie direkt mit einem Gummibärchen in Krokodilform.

Positiv ist, dass die Lehrerin eine kognitiv fordernde Anschlussaufgabe für schnellere oder besser lesende Kinder bei ihrer Unterrichtsplanung mitgedacht hat, auch wenn sie diese den Kindern beim Arbeitsauftrag noch nicht kommuniziert hat. So können die Kinder jederzeit einen schwierigeren oder leichteren Text erhalten, um eine möglichst hohe Passung ihres Fähigkeitslevels mit der Aufgabe zu erzielen.

Negativ sind hier allerdings auch zwei Dinge anzumerken. Zum einen prüft die Lehrkraft das Leseverständnis der Schülerin nicht. Sie hätte zum Beispiel sich vorlesen lassen können oder Verständnisfragen zum Text stellen können. Stattdessen vertraut sie darauf, dass die Schülerin sich einschätzen kann, lobt sie und gibt ihr sofort die Anschlussaufgabe und eine Belohnung. Und diese Belohnung ist auch direkt der zweite Kritikpunkt. Denn insgesamt stellt sich die Frage, ob bei dieser Leseübung überhaupt ein Belohnungssystem notwendig ist oder ob dadurch eventuell auch die intrinsische Motivation, die aus dem Interesse am Text und der Freude am Lesen an sich entstehen kann, auf eine extrinsische Motivation ablenkt wird, das heißt, auf den Wunsch, ein Gummibärchen zu erhalten. Damit wäre eine Belohnung im Grunde kontraproduktiv.

Die Studierenden konnten bei jeder Modellantwort wählen, ob sie diese als Audio, über dessen Transkript oder über eine Kombination der beiden Modalitäten rezipieren, und gaben jeweils anhand von vier vorstrukturierten Antwortmöglichkeiten (Abbildung 1) an, ob und mithilfe welcher Modalität sie die jeweilige Modellantwort rezipiert haben.

Nach der zweiten Videoanalyse gaben die Studierenden abschließend an, welche Modalität sie rückblickend und hinsichtlich zukünftiger Unterrichtsvideoanalysen präferieren, und begründeten ihre Präferenz mit einem offenen Kommentar. Folglich liegt bei der abschließenden Präferenzabfrage, die im Fokus dieses Beitrags steht, ein paralleler Mixed-Methods-Ansatz vor, bei dem die quantitativen Ergebnisse (Abfrage der Modalitätspräferenz) und die qualitativen Ergebnisse (Begründung der Modalitätspräferenz) komplementär in Beziehung zueinander gesetzt werden, sich ergänzen und einander näher ergründen sollen (Kelle, 2022).

Die studentischen Angaben zur bevorzugten Modalität der Modellantworten während (= Zwischennutzung) und nach der Videoanalyse (= Endpräferenz) wurden deskriptivstatistisch und varianzanalytisch ausgewertet. Die Auswertung der offenen Kommentare der Studierenden erfolgte anhand einer induktiven, strukturierenden qualitativen Inhaltsanalyse (Kuckartz & Rädiker, 2022). Die Güte des Vorgehens wurde konsensuell und argumentativ mithilfe der Berechnung einer Inter-coderübereinstimmung zwischen zwei Kodierer:innen und deren prozentualer Übereinstimmung (= PÜ) abgesichert ($0.64 < k < 1.00$; $91.2\% < PÜ < 100,0\%$).

5 Ergebnisse

Während der Intervention nutzt eine deutliche Mehrheit der Studierenden nach Selbstaussage für die Rezeption der Modellantworten das Audio ($60.6 < \% < 70.8$). In Abbildung 1a ist zu erkennen, dass das Transkript ($13.9 < \% < 24.8$) oder die Kombination aus Audio und Transkript ($10.9 < \% < 18.2$), mit Ausnahme der Modellantworten (M) 5–7, etwa gleich häufig genutzt werden. Kaum Studierende geben an, einzelne Modellantworten nicht rezipiert zu haben ($0.0 < \% < 2.9$). Dabei bleibt die Modalitätswahl während der Intervention weitgehend stabil (F1).

Etwa 62,0 % der Studierenden nutzen stets dieselbe Modalität und geben diese auch als Endpräferenz an. So ergibt sich ein starker signifikanter Zusammenhang zwischen Zwischennutzung und Endpräferenz hinsichtlich der Rezeption der Modellantworten über Audio (*Cramer's V* = 0.65; $p < .001$) oder Transkript (*Cramer's V* = 0.62; $p < .001$), nicht aber hinsichtlich der Rezeption über eine Kombination der beiden Modalitäten (*Cramer's V* = 0.37; $p < .001$). Auffällig ist, dass in allen vier Gruppen gleichermaßen deutlich mehr Studierende als Endpräferenz die Audio-Transkript-Kombination wählen (Abbildung 1c-e). Das Audio bleibt mit insgesamt 49,6 % zwar die jeweils präferierte Modalität hinsichtlich zukünftiger Videoanalysen, jedoch scheint eine Audio-Transkript-Kombination mit nunmehr 34,3 % Befürworter:innen für die Studierenden durchaus ebenfalls attraktiv zu sein (Abbildung 1b) (F2, F3).

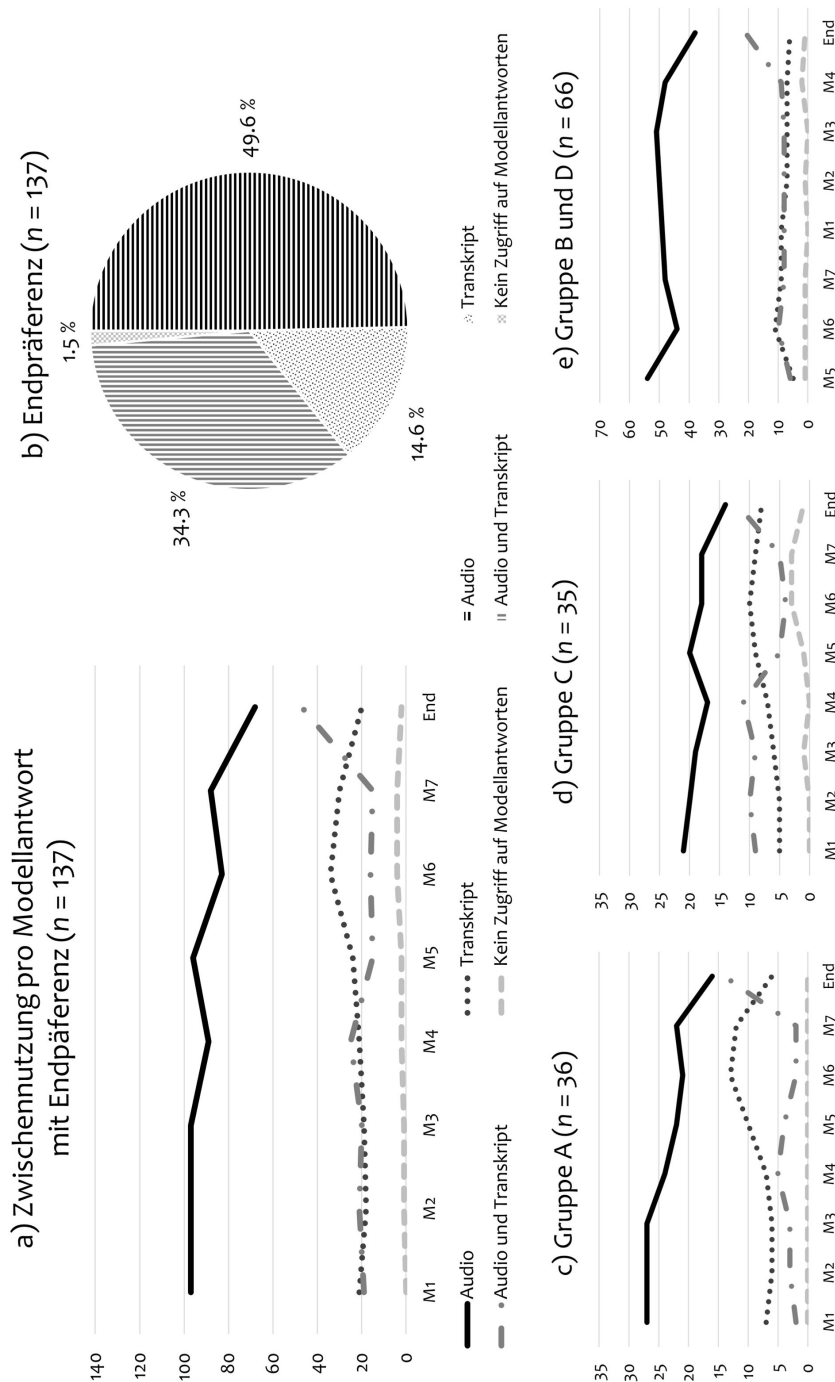


Abbildung 1: Häufigkeitsdarstellung der Modalitätswahl während und nach der Intervention in der Gesamtstichprobe und in den Interventionsgruppen (M=Modellantwort; End=Endpräferenz)

Es stellt sich die Frage, warum Studierende die Modalität der Modellantworten während der Intervention wechseln bzw. warum einige letztlich ein kombiniertes Angebot bevorzugen. Zur Beantwortung der Frage wurde der Zusammenhang zwischen den Merkmalen der Interventionsgruppen und der Zwischennutzung geprüft. Es ergaben sich insgesamt keine signifikanten Zusammenhänge zwischen den vier Ausprägungen bei der Modalitätswahl (Audio/Transkript/Kombination/kein Zugriff) und der Art der Videoanalyse (offen/fragengeleitet) bzw. der Art des Videobeispiels (funktional/dysfunktional) ($0.13 < Cramer's V < 0.32$; $.051 < p < .923$). Einzig bei der Korrelation des Transkripts als Rezeptionsmodalität und der Art des Videobeispiels zeigte sich ein schwacher, aber signifikanter Zusammenhang ($Cramer's V = 0.29$; $p < .05$). Dieser Zusammenhang erklärt sich jedoch weniger durch die Art des Videobeispiels (Abbildung 1c \triangleq d) als durch die Reihenfolge der Darbietung der Modellantworten (Abbildung 1c, d vs. e). So weichen nur die Gruppen A und C, die M5–7 im Rahmen der zweiten und nicht ersten Videoanalyse rezipiert haben, bei den besagten Modellantworten tendenziell eher auf das Transkript aus. Bei den Gruppen B und D, die M 5–7 im Rahmen der ersten Videoanalyse rezipiert haben, ist dieser Ausweich-Effekt dagegen nur leicht bei M6 zu verzeichnen. Eine mögliche Erklärung bietet die Länge der Modellantworten. Während die Audio-Dateien von M1–4 zwischen 00:53 und 02:07 Minuten dauern, umfassen M5–7 zwischen 02:13 und 05:07 Minuten. Dabei ist M6 mit 05:07 Minuten die längste Audio-Datei, wodurch sich der leichte Ausweich-Effekt der Gruppen B und D erklären lässt (F3).

Ihre Modalitätspräferenz begründeten 130 der insgesamt 137 Studierenden mit einem offenen Kommentar, fünf enthielten sich und zwei bezogen sich nicht auf die Modalität, sondern auf allgemeine Vorteile von Modellantworten (z. B. „Die Audiodatei ist immer sehr hilfreich für mich, da es manchmal Aspekte gibt, denen ich vorher nicht viel Aufmerksamkeit gegeben habe. Es ist aber auch wichtig für mich zu wissen, ob meine eigene Einschätzung korrekt ist oder ob ich mit meinen Ansichten falsch liege.“). Im Rahmen der qualitativen Inhaltsanalyse wurden die offenen Begründungen der Studierenden anhand von insgesamt 14 Kategorien ausgewertet mit dem Ziel, zu erfassen, welche Argumente für oder gegen eine Modalität genannt wurden. In der nachfolgenden Zusammenfassung wird sich auf eine Gegenüberstellung der Argumente beschränkt, mit denen die Studierenden für eine Modalitätsform argumentiert haben. Zum Vergleich wird hierfür jeweils in Klammern die Anzahl der Nennungen eines Arguments angegeben, mit denen die Präferenz von *Audio* oder *Transkript* begründet wurden (Audio:Transkript). So lässt sich vereinfacht zusammenfassen, dass die Rezeption der Audios von den Studierenden tendenziell als angenehmer (30:5), persönlicher (7:0) und, teils durch die Intonation der Dozentin, als inhaltlich klarer (11:3) empfunden wurde, konzentrierter (7:2) und weniger kognitiv belastend (13:4) erfolge, eher das Mitdenken anrege (24:6) und eine willkommene Abwechslung zur üblichen Rezeptionsform *Lesen* sei (6:1) (z. B. „Beim Lesen verliere ich sehr, sehr schnell die Konzentration. Durch das Anhören geht es schneller und ich kann besser darüber nachdenken, so als wäre ich gerade in einem Gespräch. Es fühlt sich auch weniger anstrengend und gezwungen an.“). Zudem wurde das Lernen anhand der Audios als nachhaltiger empfunden („Etwas gesagt zu bekommen, ist viel einprägsamer als nur das reine Lesen.“) (11:4). Dagegen erfolge die Rezeption des Transkripts schneller und in eigenem Tempo (8:17) und ermögliche das Notieren oder Markieren wichtiger Inhalte (1:3) sowie ein langfristiges Abspeichern für späteres Nachlesen (0:19) (z. B. „Das Transkript kann man sich gut abspeichern, bzw. abfotografieren und es weiterhin nutzen und anschauen“). Allerdings wurde auch angemerkt, dass zum einen die gewählte Modalität durch die andere Modalität ergänzt werden könne (18:19) („Falls etwas bei der Audio-Datei unklar war oder es zu schnell ging, kann man dies im Transkript noch einmal nachlesen“) und zum anderen die Wahl der Modalität auch vom Kontext, d. h. von Lernort, Technik und Tagesform abhängen könne (11:9).

Letztlich argumentierten von den 130 Studierenden, die einen offenen Kommentar abgegeben haben, 113 Studierende für das Audio und 68 Studierende für das Transkript, wobei davon insgesamt 51 Studierende für beide Modalitäten Argumente einbrachten. Unter den 51 Studierenden sind zudem auch sechs Studierende, die als Modalitätspräferenz einer der beiden Modalitäten den Vorzug gaben, in ihrer Begründung aber über die Vorteile einer Kombination nachdachten. Allerdings begründeten auch zwei Studierende, die eine Kombination präferierten, ihre Präferenz nur mit Vortei-

len von einer Modalität und eine Person, die keinen Zugriff auf die Modellantworten präferierte, erklärte: „ *fand beide Möglichkeiten in Kombination am besten*“. Abgesehen von diesen neun Studierenden entsprachen die angegebenen Begründungen der jeweiligen Modalitätspräferenz der Studierenden und deuten somit an, dass die Studierenden weitgehend eindeutig in ihrer Präferenzwahl ausgerichtet waren. Nur eine Person gab explizit an, auf Audios zugunsten der Transkript-Dateien verzichten zu können, während fünf Personen angaben, auf die Transkript-Dateien zugunsten der Audios verzichten zu können („*Das Audio ist verständlich genug, das Transkript bietet mir keinen Mehrwert.*“) (F4).

6 Diskussion

Zusammenfassend lässt sich festhalten, dass das Audio während und nach der Intervention von der Mehrheit der Studierenden für die Rezeption der Modellantworten bevorzugt wurde (F1, F2) und insgesamt als angenehmer und lernwirksamer als das Transkript beschrieben wird, während bei der Rezeption mithilfe des Transkripts eher praktische Vorteile wie die Speicherbarkeit genannt werden (F4). Das Transkript wurde anstelle des Audios bzw. einer Kombination aus beiden Modalitäten bevorzugt, wenn Modellantworten eine gewisse Länge überschritten. Dieser Ausweich-Effekt auf das Transkript wird verstärkt, wenn die Modellantwort zu einem späteren Zeitpunkt in der Intervention angeboten wurde (F3). Die Kombination der beiden Modalitäten Audio und Transkript wurde rückblickend und hinsichtlich zukünftiger Unterrichtsvideoanalysen deutlich stärker präferiert, als sie während der Intervention genutzt wurde (F3).

Insgesamt bestätigen die Ergebnisse Befunde der bisherigen Forschung zur Feedbackmodalität im Bereich Modellantworten (zsf. Alharbi, 2021). Mit Rückgriff auf die *Cognitive Theory of Multimedia* (Mayer, 2022) kann der Ausweich-Effekt vom Audio auf das Transkript anhand des sogenannten *Flüchtigkeitsprinzips* (*Transient Information Principle*) erklärt werden (Jiang & Sweller, 2022). Dieses besagt, dass Inhalte, die einen gewissen Grad an Komplexität oder Umfang überschreiten, einen anderen Anspruch an die kognitive Verarbeitung stellen und nicht auditiv-flüchtig, sondern eher in einer visuell-permanenten Form dargeboten werden sollten, um kognitiv nicht zu überlasten. Da bei der zweiten Videoanalyse die Studierenden womöglich bereits kognitiv erschöpft waren, scheint der Flüchtigkeitseffekt hier stärker als am Anfang der Intervention einzutreten. Offen bleibt die Frage, ab welcher Länge und Komplexität der Modellantworten der Effekt konkret eintritt.

Zudem bleibt unklar, warum das Audio weitgehend bevorzugt wird bzw. warum die Modalitätswahl insgesamt weitgehend stabil erscheint. Hierfür liefert die qualitative Analyse der Studierendenbegründungen erste Erklärungen. Für eine bessere Vergleichbarkeit und kontrastierende Gewichtung der hier zusammengefassten Argumente sollten die Präferenzbegründungen der Studierenden zukünftig durch eine geschlossene Abfrage überprüft werden. Eine entsprechende Abfrage mit einer neuen Studierendenkohorte wurde im Sommersemester 2025 durchgeführt und wird aktuell ausgewertet. Ebenfalls einschränkend muss eingewandt werden, dass die vorliegende Studie nicht die Lernwirksamkeit der Modalitäten geprüft hat, sondern Studierendenpräferenzen fokussiert. Die Lernwirksamkeit der Modalitäten sollte in kommenden Studien vertiefend berücksichtigt und überprüft werden.

Letztlich ermutigen die Ergebnisse, in der Hochschullehre gezielt stärker auditive Elemente zu integrieren, und sensibilisieren zugleich für die Grenzen dieser Modalität. Angesichts der Stabilität der Modalitätswahl kann es sich zudem insbesondere im Diskurs um eine inklusive Hochschule anbieten, Wahlmöglichkeiten bei der Modalität von Inhalten zu ermöglichen.

Literatur

- Alharbi, M. A. (2021). Impact of teacher written vs. Audio feedback on EFL undergraduates' writing. *Cypriot Journal of Educational Sciences*, 16(3), 1141–1153. <https://doi.org/10.18844/cjes.v16i3.5836>
- Blömeke, S. (2025). Intelligence, knowledge, skills, behavior. In A. Gegenfurtner & R. Stahnke (Hrsg.), *Teacher Professional Vision: Theoretical and Methodological Advances* (S. 57–70). Routledge. <https://doi.org/10.4324/9781003370901-7>
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies: Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- BMBF. Bundesministerium für Bildung und Forschung (Hrsg.) (2019). *Verzahnung von Theorie und Praxis im Lehramtsstudium. Erkenntnisse aus Projekten der „Qualitätsoffensive Lehrerbildung“*. Zarbock.
- Böhnke, A., Jordan, A.-M., Großmann, L., Haase, S., Helbig, K., Müller, J., Achour, S., Krüger, D. & Thiel, F. (2022). Das FOCUS-Videoportal der Freien Universität Berlin. Videobasierte Lerngelegenheiten für die erste und zweite Phase der Lehrkräftebildung. In R. Junker, V. Zucker, M. Oellers, T. Rauterberg, S. Konjer, N. Meschede & M. Holodynski (Hrsg.), *Lehren und Forschen mit Videos in der Lehrkräftebildung* (S. 37–55). Waxmann.
- Castro-Alonso, J. C. & Sweller, J. (2022). The Modality Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 261–267). Cambridge University Press.
- Delere, M. & Wilkens, L. (2024). Cognitive Load in der Arbeit mit barrierefreien Unterrichtsvideos – Audio-Deskription als Signalgeber. In V. Heitplatz & L. Wilkens (Hrsg.), *Die Rehabilitationstechnologie im Wandel: Eine Mensch-Technik-Umwelt Betrachtung* (S. 415–433). Eldorado.
- Gaudin, C. & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, 16, 41–67. <http://dx.doi.org/10.1016/j.edurev.2015.06.001>
- Holodynski, M., Meschede, N., Junker, R. & Zucker, V. (2022). Lehren und Forschen mit Videos in der Lehrkräftebildung. Eine Einführung. In R. Junker, V. Zucker, M. Oellers, T. Rauterberg, S. Konjer, N. Meschede & M. Holodynski (Hrsg.), *Lehren und Forschen mit Videos in der Lehrkräftebildung* (S. 7–16). Waxmann.
- Huxham, M. (2007). Fast and effective feedback: Are model answers the answer? *Assessment & Evaluation in Higher Education*, 32(6), 601–611. <https://psycnet.apa.org/doi/10.1080/02602930601116946>
- Jiang, D. & Sweller, J. (2022). The Transient Information Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 268–274). Cambridge University Press. <https://doi.org/10.1017/9781108894333.027>
- Kelle, U. (2022). Mixed Methods. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 163–177). Springer Fachmedien. https://doi.org/10.1007/978-3-658-37985-8_9
- König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X. & Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review*, 36, 1–18. <https://doi.org/10.1016/j.edurev.2022.100453>
- Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung: Grundlagentexte Methoden* (5. Aufl.). Beltz Juventa.
- Lotz, M. & Corvacho del Toro, I. (2013). Die Videostudie im Fach Deutsch: „Lucy rettet Mama Kroko“. In M. Lotz, F. Lipowsky & G. Faust (Hrsg.), *Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE)*. Band 3: Technischer Bericht zu den PERLE-Videostudien (S. 29–36). GFFP.
- Mayer, R. E. (2022). Cognitive Theory of Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 57–72). Cambridge University Press. <https://doi.org/10.1017/9781108894333.008>
- Renkl, A. (2022). The Worked Example Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 231–240). Cambridge University Press. <https://doi.org/10.1017/9781108894333.023>
- Seidel, T. (2022). Professionelle Unterrichtswahrnehmung als Teil von Expertise im Lehrberuf. Weiterentwicklungsperspektiven für die videobasierte Lehrerforschung. In R. Junker, V. Zucker, M. Oellers, T. Rauterberg, S. Konjer, N. Meschede & M. Holodynski (Hrsg.), *Lehren und Forschen mit Videos in der Lehrkräftebildung* (S. 17–36). Waxmann.

Zmiskol, T. & Hess, M. (2024). „Die Audio-Feedbacks waren eine totale Bereicherung.“ – Direktes Feedback bei Unterrichtsvideoanalysen in der asynchronen Online-Lehre. In A. Flügel, A. Gruhn, I. Landrock, J. Lange, B. Müller-Naendrup, J. Wiesemann, P. Büker & A. Rank (Hrsg.), *Grundschulforschung meets Kindheitsforschung reloaded* (S. 519–525). Verlag Julius Klinkhardt. <https://doi.org/10.25656/01:31411>

Autorinnen

Tabea Zmiskol. Otto-Friedrich-Universität Bamberg, Lehrstuhl für Grundschulpädagogik und -didaktik, Bamberg, Deutschland; E-Mail: tabea.zmiskol@uni-bamberg.de

Prof. Dr. Miriam Hess. Otto-Friedrich-Universität Bamberg, Lehrstuhl für Grundschulpädagogik und Grundschuldidaktik, Bamberg, Deutschland; E-Mail: miriam.hess@uni-bamberg.de



Zitiervorschlag: Zmiskol, T. & Hess, M. (2026). Multimodale Modellantworten als direktes Feedback bei der individuellen Unterrichtsvideoanalyse. *die hochschullehre*, Jahrgang 12/2026. DOI: 10.3278/HSL2614W. Online unter: wbv.de/die-hochschullehre