

Empirische Bildungsforschung trifft Geschichtsdidaktik:  
Eine Untersuchung der Lernprozesse im  
Lernarrangement mit Zeitzeug:innenberichten

Dissertation  
zur Erlangung des Doktorgrades  
der Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen

vorgelegt von

Lisa Zachrich

Tübingen

2022

1. Betreuerin:	Jun.-Prof. Dr. Christiane Bertram
2. Betreuer:	Prof. Dr. Ulrich Trautwein
Tag der mündlichen Prüfung:	26.10.2022
Dekan:	Prof. Dr. Ansgar Thiel
1. Gutachter:	Prof. Dr. Ulrich Trautwein
2. Gutachter:	Prof. Dr. Stephan Schwan
3. Gutachterin:	Jun.-Prof. Dr. Christiane Bertram

## Danksagung

Ich bin sehr dankbar für die vielen besonderen Menschen, die mich während der vergangenen Jahre beruflich sowie privat begleitet haben.

Mein größter Dank gilt meinem Betreuer:innen-Team: Jun.-Prof. Dr. Christiane Bertram, Prof. Dr. Ulrich Trautwein und Dr. Wolfgang Wagner. Vielen Dank für die intensive Zusammenarbeit, eure Ideen und die Unterstützung auf inhaltlicher, methodischer und motivationaler Ebene. Christiane, deine Begeisterung für das Lernen mit Zeitzeug:innen war seit unserem Kennenlernen ansteckend für mich. Ulrich, vielen Dank für die Freiheiten in meiner Arbeit und dein Vertrauen. Danke Wolfgang für deine immer offen stehende Tür. Ich habe so viel von euch gelernt!

Für die wertvollen Anregungen zu Beginn meiner Promotionszeit und die Zusammenarbeit möchte ich mich bei Christine Baron und Allison Weller bedanken. Ebenso bedanke ich mich bei den Mitarbeiter:innen des Koordinierenden Zeitzeugenbüros in Berlin, die mich bei der Akquise meiner Stichprobe wesentlich unterstützt haben. Ein großer Dank gilt meinen (auch ehemaligen) Kolleg:innen und Freund:innen am Hector-Institut für Empirische Bildungsforschung und der LEAD Graduate School für die offene Atmosphäre, die Hilfsbereitschaft und die vielen schönen und interessanten Begegnungen. Ganz besonders dankbar bin ich Ann-Kathrin Jaekel, Lisa Hasenbein und Moritz Fleischmann: Vielen Dank dafür, dass ich mit jeder Frage zu euch kommen konnte, für euer kritisches und ehrliches Feedback, dafür, dass ihr mir Mut zugesprochen und mich auf den letzten Metern motiviert habt. Vielen Dank an alle wissenschaftlichen Hilfskräfte des History-Teams, insbesondere Katharina Totter und Hannah Herbst, die sogar während meiner Elternzeit dieses Projekt vorangetrieben haben.

Danke Magdalena Novak für die enge Begleitung durch private und berufliche Höhen und Tiefen während der gesamten Dissertationsphase, für dein wertvolles Feedback, die emotionale Unterstützung auf langen Durststrecken und für das Mitfeiern kleiner Erfolge. Ich danke Tatjana und Bernd, die mir durch die liebevolle Betreuung meines Sohnes den Raum geschaffen haben, an einzelnen Tagen intensiv in die Arbeit einzutauchen. Zutiefst dankbar bin ich meinen Eltern: Danke für eure bedingungslose Liebe und Unterstützung in allen Lebenslagen, für euer Interesse und eure Begeisterung für mein Tun. Ihr habt damit wesentlich dazu beigetragen, dass ich dieses Promotionsprojekt glücklich gemeistert habe. Danke Tobi, für deine Unterstützung durch konstruktive Ratschläge und dein liebevolles Mitfühlen, wann immer ich das ein oder andere brauchte. Danke Michel, dass du mich immer wieder zurückgeholt hast in das bunte Leben jenseits des Schreibtisches. Danke!



---

## Zusammenfassung

Im Geschichtsunterricht sollen Schüler:innen lernen, historisch zu denken. Kompetenzen historischen Denkens umfassen unter anderem die Fähigkeiten, historische Narrationen zu de- und rekonstruieren, verschiedene Quellengattungen voneinander zu unterscheiden und in der Vergangenheit Orientierung für Gegenwart und Zukunft zu finden (z. B. Schreiber et al., 2006). Damit leistet der Geschichtsunterricht einen wesentlichen Beitrag zu dem Bildungsziel der Schulen, Schüler:innen zu eigenständigem Urteilen zu befähigen (z. B. Baden-Württemberg: Schulgesetz § 1, Absatz 2). Neben dieser gesellschaftlichen Relevanz des Geschichtsunterrichts hat er aufgrund des Anspruchs, Orientierung zu stiften, auch eine persönliche Relevanz für die Lernenden. Damit geht einher, dass den Emotionen der Lernenden im Geschichtsunterricht eine besondere Bedeutung zukommt, denn sie sind Teil der Erkenntnisprozesse für das persönliche Leben (Rüsen, 2020). Das Lernen mit Zeitzeug:innenberichten ist ein beliebtes Lernarrangement im Geschichtsunterricht (Henke-Bockschatz, 2007, 2014; Whitman, 2004), in dem die Verschränkung von kognitiven und emotionalen Lernprozessen sichtbar wird: Denn die persönlichen Geschichten erzählen nicht selten von persönlichen Herausforderungen und machen Orientierungs- und Identifikationsangebote. Allerdings zeigen empirische Studien, dass die persönlichen Geschichten die Gefahr bergen, Schüler:innen die Fehlvorstellung von Geschichte zu vermitteln, die Vergangenheit sei über eine einzige Perspektive zu verstehen (Ballis & Schwendemann, 2021; Bertram et al., 2017; Brüning, 2018; Hogervorst, 2020; Obens & Geißler-Jagodzinski, 2010; Richardson, 2021). Mit diesem Konflikt zwischen den Zielen des Geschichtsunterrichts und der Wirkung von Zeitzeug:innenberichten eröffnen sich zentrale Fragen zum Verständnis der Lernprozesse in diesem Lernarrangement. Zum Beispiel: Wie nehmen Lernende die Berichte wahr? Wie gestalten sich die Lernprozesse der Lernenden? Welche Rolle spielen dabei Emotionen?

Geschichtsdidaktische Forschung widmete sich bisher kaum der systematischen Erforschung der Lernprozesse in diesem Lernarrangement, sondern nahm vorwiegend die Wirkungen der Berichte in den Blick (z. B. Bertram et al., 2017). Das bedeutet, dass es an einem umfassenden Verständnis der Lernprozesse fehlt. Die Empirische Bildungsforschung, konkret die allgemeine Unterrichtsforschung, kann Anknüpfungspunkte für diese systematische und empirische Erforschung der Lernprozesse bieten. Sie hat sich bisher allerdings nur selten mit Fächern der historisch-politischen Bildung beschäftigt (z. B. Trautwein et al., 2021), sondern den Fokus auf mathematisch-naturwissenschaftliche Fächer sowie die allgemeinen Lesefähigkeiten gelegt (z. B. Maag Merki, 2021; Seidel & Shavelson, 2007). Ziel der vorliegenden Dissertation ist es,

durch die Integration der Perspektiven von Geschichtsdidaktik und Empirischer Bildungsforschung ein systematisches Verständnis für die Lernprozesse beim Lernen mit Zeitzeug:innenberichten zu bekommen und aus der interdisziplinären Perspektive Erkenntnisse für Geschichtsdidaktik und Empirische Bildungsforschung abzuleiten.

Das Angebots-Nutzungs-Modell zur Beschreibung unterrichtlicher Wirkungsweisen (Helmke, 2017) dient in der vorliegenden Dissertation als Rahmenmodell. Das Modell identifiziert wichtige Einflussfaktoren im komplexen Unterrichtsgeschehen und wird in dieser Arbeit durch das Lernarrangement mit Zeitzeug:innenberichten mit fachspezifischen Inhalten gefüllt. Es wird herausgearbeitet, dass insbesondere die Lernprozesse im Rahmen des Lernarrangements bisher eine Blackbox sind. Ein systematisches Verständnis über diese Prozesse erscheint allerdings vielversprechend, um die Wirkungen des Lernarrangements sowie die Verschränkung von Emotionen und Kognitionen besser zu verstehen. Die Dissertation stützt sich zur Erreichung des Forschungsziels auf eine konzeptionelle Studie sowie drei empirische Studien, die unterschiedliche Bereiche des Angebots-Nutzungs-Modells thematisieren und mit unterschiedlichen Designs erforschen.

In Studie 1 (*Historical experiences: A framework for encountering complex historical sources*) wurde ein konzeptuelles Modell zur Beschreibung der besonderen Lernerfahrung mit komplexen historischen Quellen entwickelt. Denn Grundlage für die systematische Erforschung des Lernarrangements ist ein Verständnis darüber, welche Lernprozesse hier von Bedeutung sind. Das Modell fußt auf einer Verknüpfung verschiedener wissenschaftlicher Diskussionslinien zu vergleichbaren Lernerfahrungen und identifiziert kognitive, affektive und physische Involviertheitsprozesse als Teil der besonderen Lernerfahrung mit Zeitzeug:innenberichten und historischen Orten. Die konzeptionelle Studie sollte für die Forschungsgemeinschaft einen Ausgangspunkt bieten, die Lernprozesse in diesen Lernarrangements zu erforschen und zu diskutieren.

Studie 2 (*Das „Aura-Erlebnis“ in der Begegnung mit Zeitzeug\*innen – eine komplexe Lernerfahrung verstehen und standardisiert erfassen*) ging der Frage nach, ob die Verarbeitungsprozesse mit Zeitzeug:innenberichten durch einen standardisierten Fragebogen messbar gemacht werden können. Hierfür wurden zwei Pilotierungsstudien mit Studierenden und eine Validierungsstudie mit 668 Schüler:innen durchgeführt. Das Ergebnis der Studien war die Validierung des standardisierten Fragebogens zur Erfassung der Prozesse im Rahmen der distanzlos-involvierten Verarbeitung, welcher für weitere Studien genutzt werden kann. Darüber hinaus wurden in dieser Studie Fragen zur Validität der Konstrukte beantwortet. Die Ergebnisse zeigten, dass die Elemente der besonderen Lernerfahrung in negativem Zusammenhang mit

---

einer kritisch-distanzierten Verarbeitung des Lernarrangements stehen. Außerdem zeigte sich ein Zusammenhang von individuellen Voraussetzungen der Lernenden mit den Verarbeitungsprozessen.

Studie 3 (*Really? It depends! How authentic learning material affects involvement and learning with personal stories of the past*) untersuchte in einem experimentellen Online-Setting mit 452 Teilnehmenden, welchen Effekt die Authentizität videografiertes Zeitzeug:innenberichte als ein zentrales Charakteristikum von Zeitzeug:innenberichten auf die Verarbeitungsprozesse hatte und welche Rolle die wahrgenommene Authentizität in diesem Zusammenhang spielte. Je nach Art der Einführung ein und desselben Videoclips – als originaler Bericht, als nachgespielter Bericht oder als rekonstruierter Bericht – zeigten sich jeweils unterschiedliche Effekte auf die affektiven Verarbeitungsprozesse: Lernende, die den Zeitzeug:innenbericht mit der Einführung sahen, der Clip sei ein originaler Bericht, waren stärker affektiv involviert als Lernende, die die Zeitzeug:innenberichte mit den anderen Einführungen sahen. Zudem zeigte sich, dass die wahrgenommene Authentizität in Form der Einschätzung, die Berichte machen die Vergangenheit lebendig, prädiktiv für die Verarbeitungsprozesse war.

Studie 4 (*When some students are better off with lessons „as usual“: An aptitude-treatment-interaction study in the history classroom*) folgte der Frage, inwiefern die individuellen Lernvoraussetzungen für die Wirkung des Formats des Zeitzeug:innenberichts auf das Lernen historischer Kompetenzen eine Rolle spielten. Konkret wurde anhand einer randomisiert-kontrollierten Interventionsstudie mit 769 Schüler:innen untersucht, ob Lernende mit unterschiedlichem Wohlbefinden in sozialen Situationen einen unterschiedlichen Lernzuwachs beim Lernen mit videografierten, verschriftlichten oder Live-Berichten erzielten. Es zeigte sich, dass es insbesondere für Lernende, die sich in sozialen Situationen unwohl fühlten, von Bedeutung ist, mit welchem Format des Zeitzeug:innenberichts sie lernten (Live, Video, Text), während Lernende, die sich in sozialen Situationen wohlfühlten, mit allen Formaten gleich gut lernten.

Durch die Verknüpfung von Empirischer Bildungsforschung mit Geschichtsdidaktik wird in der vorliegenden Dissertation ein umfassendes Verständnis der Lernprozesse im Lernarrangement mit Zeitzeug:innenberichten erzielt. Darüber hinaus wird im Rahmen der Dissertation der Erkenntnisgewinn der interdisziplinären Forschungszugänge für die beteiligten Disziplinen abgeleitet. Schließlich arbeitet die Dissertation Implikationen für die weitere Erforschung des Lernarrangements sowie für die Nutzung von Zeitzeug:innenberichten in der Praxis heraus.

## Inhalt

<b>1 Einleitung und theoretischer Rahmen .....</b>	<b>1</b>
1.1 Verknüpfung von Empirischer Bildungsforschung und Geschichtsdidaktik .....	6
1.1.1 Das Angebot-Nutzungs-Modell zur Beschreibung von Unterricht .....	6
1.1.2 Historisches Denken als Ziel des Geschichtsunterrichts .....	9
1.1.3 Emotionen im (Geschichts-)Unterricht .....	15
1.2 Unterrichtsangebot: Das Lernarrangement mit Zeitzeug:innenberichten.....	18
1.2.1 Zeitzeug:innenberichte als Lernmaterial: konzeptuelle Überlegungen .....	18
1.2.2 Zeitzeug:innenberichte und ihre Einbettung im Unterricht.....	21
1.2.3 Chancen und Risiken des Lernarrangements: Empirische Studien .....	25
1.2.4 Oberflächen- und Tiefenmerkmale des Lernarrangements mit Zeitzeug:innenberichten .....	29
1.3 Lernpotenzial, Wahrnehmung und Lernaktivität: Lernende begegnen Zeitzeug:innenberichten.....	32
1.3.1 Individuelle Lernvoraussetzungen und ihr Zusammenspiel mit der Unterrichtsgestaltung.....	32
1.3.2 Wahrnehmung und Interpretation: Authentizität von Zeitzeug:innenberichten .....	35
1.3.3 Lernaktivität: Vier Verarbeitungsmodi im Umgang mit Zeitzeug:innenberichten .....	39
<b>2 Zielsetzung und Forschungsfragen.....</b>	<b>46</b>
<b>3 Studie 1: Historical Experiences: A Framework for Encountering Complex     Historical Sources.....</b>	<b>51</b>
<b>4 Studie 2: Das „Aura-Erlebnis“ in der Begegnung mit Zeitzeug*innen – Eine     komplexe Lernerfahrung verstehen und standardisiert erfassen .....</b>	<b>93</b>
<b>5 Studie 3: Really? It depends! How Authentic Learning Material Affects     Involvement and Learning with Personal Stories of the Past. ....</b>	<b>117</b>
<b>6 Studie 4: When Some Students Are Better Off With Lessons “As Usual”: An     Aptitude-Treatment-Interaction Study in the History Classroom.....</b>	<b>155</b>



---

<b>7 Gesamtdiskussion</b> .....	<b>210</b>
7.1 Diskussion der Ergebnisse.....	212
7.1.1 Das Lernen mit Zeitzeug:innenberichten verstehen: Beitrag zur Geschichtsdidaktik.....	213
7.1.2 Lehr-Lernprozesse im (Fach-)Unterricht: Beitrag zur Empirischen Bildungs- und Unterrichtsforschung .....	217
7.2 Stärken und Grenzen der Studien .....	222
7.3 Implikationen für Forschung und Praxis .....	225
7.3.1 Implikationen und Anknüpfungspunkte für zukünftige Forschung .....	225
7.3.2 Implikationen für die Praxis .....	228
7.4 Fazit .....	230
<b>8 Literatur</b> .....	<b>232</b>



---

# 1

EINLEITUNG  
UND  
THEORETISCHER RAHMEN

---

# 1 Einleitung und theoretischer Rahmen

Im November 2020 demonstrierten Menschen, die sich selbst als „Querdenker“ bezeichneten, gegen die Maßnahmen der Regierung zur Eindämmung der Corona-Pandemie. Schlagzeilen machten dabei unter anderem zwei historische Vergleiche: Ein elfjähriges Mädchen verglich ihre Situation während des Lockdowns mit jener von Anne Frank, die sich während des Zweiten Weltkriegs als jüdisches Kind in Amsterdam vor den Nationalsozialisten verstecken musste, um nicht in ein Konzentrationslager deportiert und womöglich vergast zu werden. Eine Demonstrantin in Hannover verglich ihr Engagement bei den „Querdenkern“ mit dem von Sophie Scholl, die im Zweiten Weltkrieg im Widerstand gegen die Nationalsozialisten kämpfte und dafür hingerichtet wurde (Deutsche Presse Agentur [dpa], 2020).

Diese Beispiele sind Extreme. In Verbindung mit weiteren gesellschaftlichen Entwicklungen wie zunehmendem Antisemitismus auf deutschen Schulhöfen (Bernstein, 2021) und einer hohen Anzahl deutscher Jugendlicher, die im Internet nicht zwischen Meinungen und Fakten unterscheiden können – dies betrifft 45 % der 15-Jährigen (Sälzer, 2021) – verdeutlichen sie die Notwendigkeit, in der Schule Kompetenzen zu fördern, die eine reflektierte Einordnung historischer Ereignisse, eine kritische Analyse von Quellen und Darstellungen, von Fakten und Meinungen ermöglichen.

Der Geschichtsunterricht an Schulen in westlichen Demokratien hat den Anspruch, Schüler:innen<sup>1</sup> Kompetenzen historischen Denkens zu vermitteln (z. B. Hasberg, 2005; National Center for History in the Schools, 1996; Seixas & Morton, 2013). Zu diesen Kompetenzen gehören beispielsweise die Fähigkeiten, historische Ereignisse zu kontextualisieren, unterschiedliche Quellengattungen voneinander zu unterscheiden und verschiedene Quellen kritisch miteinander zu vergleichen. Die Fähigkeiten sollen dazu beitragen, dass Schüler:innen bis zum Ende ihrer Schullaufbahn ein reflexives Geschichtsbewusstsein ausbilden. Damit leistet der Geschichtsunterricht einen wesentlichen Beitrag zum Bildungsziel von Schulen, Schüler:innen dazu zu befähigen, selbstständige Urteile fällen zu können (z. B. Baden-Württemberg: Schulgesetz § 1, Absatz 2). Historisches Denken umfasst aber nicht nur den Blick in die Vergangenheit, sondern auch den Blick auf die Gegenwart und in die Zukunft: Durch den Umgang mit Geschichte sollen Lernende eine Orientierung für das persönliche gegenwärtige und zukünftige Leben finden. Mit dem Anspruch, Relevanz für das persönliche Leben der Lernenden

---

<sup>1</sup> Die Form des *Genderns* im Theoretischen Rahmen wie in der Diskussion unterscheidet sich von der Form des Genderns in Studie 2. Die unterschiedlichen Formen sind der zeitlichen Entwicklung geschuldet. Die Schreibweise mit : ist angenehmer für den Lesefluss als das \*. In beiden Fällen sind damit alle Menschen eingeschlossen, die sich zwischen oder außerhalb der binären Geschlechterzuschreibung verorten.

zu haben (z. B. Gautschi, 2009; Van Straaten, Wilschut & Oostdam, 2016), kann der Geschichtsunterricht nicht allein die kognitive Auseinandersetzung mit historischen Ereignissen und das Lernen methodischen Handwerkszeugs beinhalten. Er umfasst immer auch emotionale Komponenten – auf Ebene der historischen Akteure (Was haben die Menschen gefühlt, als die Mauer fiel?) sowie der Lernenden selbst (Was hat die deutsche Wiedervereinigung mit mir persönlich zu tun?).<sup>2</sup>

Ein spezifisches Lernarrangement im Geschichtsunterricht ist das Lernen mit Zeitzeug:innenberichten, also persönlichen Erfahrungsberichten über die Vergangenheit. Hier wird die Verschränkung von Kognition und Emotion besonders deutlich: Der Unterricht mit diesen Berichten verspricht, die Vergangenheit durch die lebendigen und persönlichen Geschichten greifbarer machen, denn „mit den Lebensgeschichten lässt sich mitfiebern und mitleiden. (...) Zeitzeugenerzählungen sind per se hochemotional“ (Brauer, 2017). Lernende werden mit den Gefühlen der Zeitzeug:innen konfrontiert, die nicht selten eine Botschaft für Gegenwart und Zukunft mitbringen und somit den Bogen zum persönlichen Leben der Schüler:innen spannen (Bertram, 2020). Gleichzeitig sollen die Schüler:innen etwas über die Vergangenheit und womöglich über den historischen Erkenntnisprozess lernen – also auch kognitiv angesprochen werden. Der Unterricht mit den individuellen Berichten kann auf Lernende motivierend wirken und ihnen die Relevanz der Vergangenheit für das persönliche Leben aufzeigen (Bertram, Wagner & Trautwein, 2017; Henke-Bockschatz, 2007). Allerdings bergen die individuellen Geschichten auch die Gefahr, den Schüler:innen die Fehlvorstellung von Geschichte zu vermitteln, dass die Vergangenheit über einen einzigen subjektiven Bericht verstanden und nacherlebt werden könne (Bertram et al., 2017; Brauer, 2020; Brauer & Lücke, 2022).

Das Lernarrangement hat folglich Potenzial, die Lernenden emotional und persönlich anzusprechen, und gleichzeitig den Anspruch, die Lernenden nicht zu überwältigen, sondern zu historischem Denken zu befähigen. Dabei ist die Frage offen, inwiefern die individuellen Geschichten den Zugang zum historischen Lernen fördern oder behindern können. Geschichts-didaktische Forschung hat sich bisher zwar mit den Chancen und Risiken des Lernarrangements befasst (z. B. Brüning, 2018; Hogervorst, 2020; von Plato, 2001), allerdings nur in Ausnahmen mit einem starken methodischen Design (Bertram et al. 2017) und bisher nicht mit dem Fokus auf Lernprozesse. Das heißt, es fehlt an einem tiefgreifenden und systematischen Verständnis über die zugrunde liegenden Lernprozesse in diesem Lernarrangement. Das macht das Lernen

---

<sup>2</sup> In der vorliegenden Dissertation werden die Begriffe Emotion, Gefühl, Affekt und die entsprechenden Adjektive synonym verwendet.

mit Zeitzeug:innenberichten zu einem besonders interessanten Forschungsfeld für die Empirische Bildungsforschung. Diese hat sich seither vorwiegend mit Bezügen zu mathematisch-naturwissenschaftlichen Fächern oder der allgemeinen Lesefähigkeit beschäftigt und nur selten Lerngelegenheiten der historisch-politischen Bildung in den Blick genommen (z. B. Maag Merki, 2021; Seidel, 2008; Seidel & Shavelson, 2007). Dabei zeigen interdisziplinäre Projekte von Bildungsforschung und Fachdidaktik den Erkenntnisgewinn der gemeinsamen Forschung auf (Praetorius, Herrmann et al., 2020; Trautwein et al., 2017; Wacker et al., 2020). In Anknüpfung an jüngste Bemühungen zur Integration von Unterrichtsforschung als Teil der Empirischen Bildungsforschung und Fachdidaktik Geschichte (Trautwein, Schreiber & Hasenbein, 2021) widme ich mich in der vorliegenden Arbeit zentralen Fragen der historischen Bildung durch den Zugang der Empirischen Bildungsforschung. Das übergeordnete Ziel der Dissertation ist es, über die systematische Untersuchung des Lernarrangements ein tiefgreifendes Verständnis für die Lernprozesse beim Lernen mit Zeitzeug:innenberichten zu bekommen. Konkreter war für die Erstellung der Dissertation das Interesse daran leitend, wie Schüler:innen in diesem Lernarrangement lernen, welche Bedeutung dabei Emotion und Kognition zukommen und welche Merkmale auf der Ebene der Unterrichtsgestaltung, der Lernenden und ihrer Wahrnehmungen eine Rolle spielen. Die Dissertation leistet damit einen wichtigen Beitrag zur Beantwortung der Frage, inwiefern die emotionale Involviertheit mit den persönlichen Erfahrungsberichten das Lernen fördern oder aber den Zugang zu historischem Denken verschließen kann (Bertram et al., 2017; Brauer & Lücke, 2013).

Hierfür nahm Studie 1 zunächst das Lernarrangement aus theoretischer Perspektive in den Blick und entwarf ein konzeptionelles Modell zur Beschreibung der Lernerfahrung mit Zeitzeug:innenberichten (und historischen Orten). Darauf aufbauend entwickelte und validierte ich in Studie 2 einen standardisierten Fragebogen zur Erfassung der situationalen Verarbeitung und Involviertheit der Lernenden mit dem Zeitzeug:innenbericht und untersuchte den Einfluss von situationalen und persönlichen Merkmalen auf die Verarbeitung der Berichte. Studie 3 untersuchte in einem experimentellen Online-Setting, welchen Effekt ein zentrales Merkmal der Berichte – die Authentizität – auf die Verarbeitung des Lernmaterials hat. Schließlich ging ich in Studie 4 mittels Sekundärdatenanalysen der Frage nach, welche Rolle das Zusammenspiel von Format der Zeitzeug:innenberichte und persönlichen Merkmalen im Lernarrangement auf den Erwerb historischer Kompetenzen hat.

Die Dissertation ist wie folgt strukturiert: Das erste Kapitel bildet den theoretischen Rahmen. Hier führe ich im ersten Teil zunächst in Grundlagen der Empirischen Bildungsforschung (Angebots-Nutzungs-Modell) und der Geschichtsdidaktik (Kompetenzen historischen

Denkens) ein und thematisiere die Rolle von Emotionen beim historischen Lernen (Kapitel 1.1). Orientierend am Angebot-Nutzungs-Modell beschreibe ich im zweiten Teil das Lernarrangement mit Zeitzeug:innenberichten als Unterrichtsangebot (Kapitel 1.2) und gehe im dritten Teil auf die Lernenden, ihre Lernvoraussetzungen, ihre Wahrnehmung des Unterrichtsangebots und ihre Lernaktivität ein (Kapitel 1.3). Im zweiten Kapitel gebe ich einen Überblick über das Ziel und die Forschungsfragen der Dissertation (Kapitel 2). Nach der Vorstellung der vier Studien (Kapitel 3–6) bietet das siebte Kapitel eine Gesamtdiskussion der zentralen Ergebnisse der Studien (Kapitel 7.1), der Stärken und Limitationen dieser Arbeit (Kapitel 7.2) sowie der Implikationen für Forschung und Praxis (Kapitel 7.3). Ich schließe die Dissertation mit einem Fazit (Kapitel 7.4).

## **1.1 Verknüpfung von Empirischer Bildungsforschung und Geschichtsdidaktik**

In diesem Kapitel werden grundlegende Modelle aus der Empirischen Bildungsforschung und Geschichtsdidaktik eingeführt: das Angebots-Nutzungs-Modell zur Beschreibung von Unterricht als Rahmenmodell dieser Dissertation sowie historische Kompetenzen als Ziel des Geschichtsunterrichts. Zudem wird die Rolle von Emotionen im Unterrichtsgeschehen, insbesondere mit Blick auf das Lernen historischen Denkens thematisiert.

### **1.1.1 Das Angebot-Nutzungs-Modell zur Beschreibung von Unterricht**

Bis zum Ende ihrer Schullaufbahn verbringen Schüler:innen in Deutschland bis zu 15.000 Unterrichtsstunden in der Schule (Fend, 2006). Durch die Etablierung der Ganztagsbetreuung sind es mittlerweile noch einige weitere Stunden. Mit Blick auf diese Zahl erscheint es wenig überraschend, dass sich die allgemeine Unterrichtsforschung als zentrales Forschungsfeld der Empirischen Bildungsforschung etabliert hat (Maag Merki, 2021). Wissenschaftler:innen gehen in diesem Teilgebiet der Empirischen Bildungsforschung beispielsweise den Fragen nach, welche Merkmale qualitätvollen Unterricht ausmachen, welche Bedeutung die Lehrkraft für die Gestaltung des Unterrichts hat und wie die Lernprozesse von Schüler:innen durch den Unterricht unterstützt werden können. Seidel und Reiss (2014) definieren Unterricht als das „Gestalten von Lernumgebungen mit dem Ziel, optimale Gelegenheiten für die effektive Ausführung von Lernaktivitäten der Schüler bereitzustellen“ (S. 254). Für die Erforschung der vielschichtigen Unterrichtsprozesse können Angebots-Nutzungs-Modelle unterrichtlicher Wirkungsweisen einen Orientierungsrahmen bieten (z. B. Creemers, 1994; Helmke, 2017; Kohler & Wacker, 2013; Praetorius, Rogh & Kleickmann, 2020; Seidel, 2014; Seidel & Shavelson, 2007; Vieluf, Praetorius, Rakoczy, Kleinknecht & Pietsch, 2020). Die Angebots-Nutzungs-Modelle stellen eine Erweiterung des aus den USA stammenden Prozess-Produkt-Paradigmas dar, das im Kern das Lernergebnis als Produkt des Lernprozesses definiert und zwischen Voraussetzungen, Prozess-, Produkt- und Kontext-Variablen unterscheidet (Dunkin & Biddle, 1974; Vieluf et al., 2020). Helmut Fend veröffentlichte 1998 eine erste deutschsprachige Version des Rahmenmodells, das fortan von vielen Wissenschaftler:innen weiterentwickelt wurde (z. B. Baumert & Klieme, 2001; Helmke, 2017; Helmke & Weinert, 1997; Seidel, 2014).<sup>3</sup> Das Angebots-Nutzungs-Modell zeigt sowohl die relevanten Strukturen und Komponenten im Un-

---

<sup>3</sup> In der englischsprachigen Forschungslandschaft entwickelten Brühwiler und Blatchford (2011) das dem Angebots-Nutzungs-Modell entsprechende Supply-Use-Model.



terrichtsgeschehen auf als auch Prozesse, wie diese Komponenten miteinander zusammenhängen können (Seidel, 2014). Es eröffnet damit viele Anknüpfungspunkte für die Erforschung des komplexen Unterrichtsgeschehens. Abbildung 1 zeigt das Modell nach Helmke (2017) in seiner jüngsten Überarbeitung.

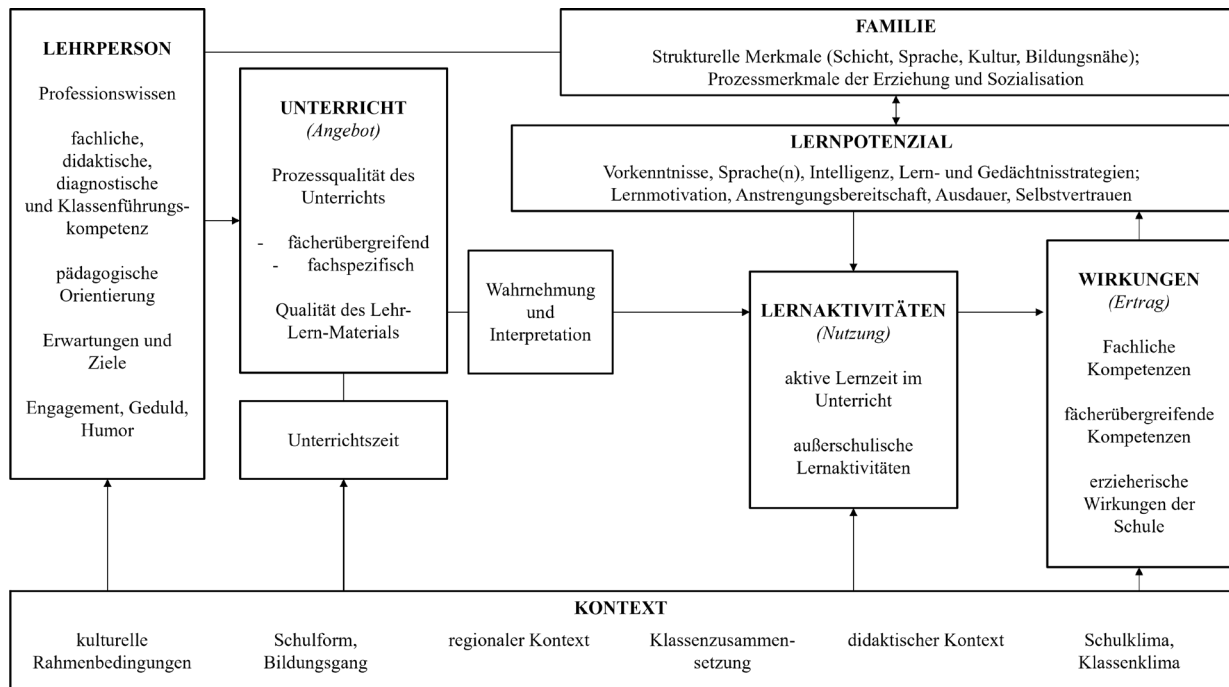


Abbildung 1. Angebots-Nutzungs-Modell der Wirkungsweise von Unterricht nach Helmke (2017)

Das Angebot-Nutzungs-Modell versteht Unterricht als Angebot, das von Schüler:innen aktiv genutzt werden muss bzw. kann, um zu dem gewünschten Ergebnis zu gelangen – allerdings sei, so Kunter und Trautwein (2013), Letzteres nicht garantiert. Das Modell unterscheidet die drei Ebenen *Angebot*, *Nutzung* und *Ertrag*. Der Unterricht, das *Angebot*, wird von der Lehrperson gestaltet. Kompetenzen, Wissen, Erwartungen, Ziele, aber auch Engagement, Geduld und Humor der Lehrperson beeinflussen, wie sie den Unterricht plant und durchführt. Helmke identifizierte die fachübergreifende und fachspezifische Prozessqualität des Unterrichts, die Qualität des Lehr-Lern-Materials sowie die Unterrichtszeit als Merkmale des unterrichtlichen Angebots. Mit dem Merkmal „Wahrnehmung und Interpretation“ wird das Modell der Annahme gerecht, dass Schüler:innen den Unterricht und seine Inhalte abhängig von ihrem Vorwissen und ihrer Vorerfahrung wahrnehmen und interpretieren: „[...] es wird eine subjektiv geprägte Interpretation konstruiert“ (Vieluf et al., 2020, S. 72). Vieluf et al. (2020) wiesen darauf hin, dass ein Bewusstsein dieser subjektiven Wahrnehmung des Unterrichts bei dem Verständnis darüber hilft, weshalb die Durchführung des Unterrichts von der Planung abweichen kann und „Kommunikation im Unterricht manchmal misslingt“ (S. 72). Wie das Angebot

genutzt wird, hängt nun von den Schüler:innen ab. Unter der *Nutzung* identifizierte Helmke in seinem jüngsten Modell die aktive Lernzeit im Unterricht sowie die außerschulischen Aktivitäten wie beispielsweise Hausaufgaben. Dabei sind genau hier die Lern- und Verarbeitungsprozesse zu verorten (Kunter & Trautwein, 2013), d. h., hier findet Lernen statt. Einfluss auf die Lernaktivität hat das Lernpotenzial der Schüler:innen. Die Lernenden unterscheiden sich beispielsweise in ihrer Intelligenz, in ihren sprachlichen Fähigkeiten und ihrer Lernmotivation. Dieses Lernpotenzial steht in engem Zusammenhang zur Familie der Schüler:innen, denn diese unterscheiden sich in strukturellen Merkmalen und der Erziehungs- und Sozialisationskultur. Die Lernaktivität, also die *Nutzung* des unterrichtlichen Angebots, soll dann im Idealfall den Kompetenzzuwachs erzielen, also Wirkungen bzw. *Erträge* zeigen. Helmke verdeutlichte mit der Berücksichtigung der Kontextmerkmale, dass die Dynamiken im Unterricht, die Lehr-Lernprozesse, immer in einen bestimmten Kontext eingebettet sind – in das Schulklima, eine bestimmte Schulform, den kulturellen oder regionalen Kontext.

Seidel (2014) wies darauf hin, dass eine Schärfung des generischen Angebots-Nutzungs-Modells durch die Zusammenarbeit von Pädagogik, Psychologie und insbesondere auch der Fachdidaktiken wünschenswert sei. Im Bereich der Fachdidaktik Geschichte orientierten sich bisher vereinzelte theoretische Überlegungen an dem Angebots-Nutzungs-Modell zur Beschreibung des Unterrichts. Gautschi (2009) entwickelte beispielsweise in seiner empirischen Studie zu gutem Geschichtsunterricht ein Rahmenmodell, welches den Unterricht durch Angebots- und Nutzungsprozesse beschreibt. In der Theorie des Geschichtsunterrichts (Bracke et al., 2018; Köster, Thünemann & Zülsdorf-Kersting, 2017) skizzierten die Autor:innen den Geschichtsunterricht als soziales System zwischen Angebot und Nutzung und konkretisierten das Verhältnis zwischen Angebot und Nutzung durch die systemtheoretische Perspektive. Zwar sei eine theoretische Trennung von Angebot und Nutzung möglich, in der Praxis, so Thünemann (2021), ließen sich die beiden Strukturen allerdings kaum voneinander trennen. Auch im Handbuch *Wirksamer Geschichtsunterricht* (Kuchler & Sommer, 2018) führten Wilhelm, Rehm und Reinhardt (2018) im Einführungskapitel das Angebots-Nutzungs-Modell als Orientierungsrahmen ein. Diese vereinzelt Verknüpfungen von Unterrichtsforschung und Fachdidaktik stellen Bemühungen dar, das generische Modell mittels der Perspektive der Fachdidaktik inhaltlich auszudifferenzieren und zu konkretisieren. Allerdings verbleiben bisherige Forschungsprojekte häufig auf theoretischer Ebene (mit Ausnahmen, z. B. Gautschi, 2009). Im folgenden Kapitel werde ich das Ziel des Geschichtsunterrichts, historisches Denken lernen, als Ertrag im Angebots-Nutzungs-Modell verorten und ausführen.

### 1.1.2 Historisches Denken als Ziel des Geschichtsunterrichts

„Geschichte hilft ihnen (den Lernenden; Anm. d. V.) [...], die Welt der Gegenwart besser zu verstehen und Orientierung für die Gestaltung ihrer Zukunft zu gewinnen“ (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016). Ein Blick in die Bildungs-, Kern- und Rahmenlehrpläne der Bundesländer zum Fach Geschichte zeigt, wie vielschichtig die Zielsetzungen des Geschichtsunterrichts sind: Schüler:innen sollen nicht nur etwas über die Vergangenheit erfahren, sondern zu einer kritischen Teilhabe an der Gesellschaft befähigt werden, sich eigene Urteile bilden und sich in der Gegenwart orientieren können. All das äußert sich in einem reflexiven Geschichtsbewusstsein der Schüler:innen (z. B. Landesinstitut für Schule und Medien Berlin-Brandenburg, 2015; Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016; Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014). Das Geschichtsbewusstsein, das als Konstrukt der deutschsprachigen Geschichtsdidaktik insbesondere von Jörn Rüsen (1994), Karl-Ernst Jeismann (1977), Hans-Jürgen Pandel (1987) und Bodo von Borries (1988) seit den 1970er Jahren theoretisch geprägt und empirisch geprüft wurde, definierte Pandel (2009) als einen „psychischen Verarbeitungsmodus historischen Wissens“ (S. 80). Für ein reflexives Geschichtsbewusstsein ist die Fähigkeit zentral, eine Verbindung zwischen Vergangenheit, Gegenwart und Zukunft herzustellen (z. B. Galda, 2013). Diese Verknüpfung vollzieht sich im Prozess historischen Denkens – und das Lernen historischen Denkens, so Schreiber (2018), ist das „Herzstück des Geschichtsunterrichts“ (S. 149).

#### *Kompetenzen historischen Denkens*

Mit der Einführung der Kompetenzdebatte nach den Ergebnissen der PISA-Studie (Weinert, 2001) bemühte sich auch die Geschichtsdidaktik um Kompetenzmodelle zur Beschreibung der Fähigkeiten, die historisches Denken ausmachen und somit ein reflexives Geschichtsbewusstsein abbilden können. Derzeit existieren mehrere Kompetenzmodelle historischen Denkens parallel, „die sich in ihren Grundgedanken und Zielsetzungen allerdings sehr ähnlich sind“ (Bertram, 2017, S. 56) und unterschiedlichen Geschichtslehrplänen zugrunde liegen. Zwar nutzen die Autor:innen der Modelle unterschiedliche Begrifflichkeiten, aber alle bilden beispielsweise die Operationen ab, auf Basis von Quellen eigene Narrationen zu entwickeln und vorhandene historische Narrationen kritisch zu analysieren (z. B. Bertram, 2017; Gautschi, Hodel, Utz & August, 2009; Pandel, 2005; Schreiber et al., 2006). In Abgrenzung zur deutschsprachigen Kompetenzdebatte lassen sich in der internationalen Forschungslandschaft weitere Traditionslinien identifizieren, die das Theorieverständnis zum historischen

Denken geprägt haben (siehe für einen Überblick: Bracke et al., 2018; Trautwein et al., 2017). Ihnen ist die Intention gemein, dass Geschichtsunterricht den Schüler:innen vermitteln soll, historisch zu denken: Im US-amerikanischen Raum erarbeitete Wineburg (1991) drei Strategien, mit denen professionelle Historiker:innen im Vergleich zu Noviz:innen historische Texte erschlossen: (1) *Sourcing*, das Erkennen der Quelle und der Absicht des Textes; (2) *Contextualization*, die Einordnung des Textes in seinen zeitlichen und räumlichen Kontext; (3) *Corroboration*, der Vergleich mehrerer Quellen. Historisches Denken von Schüler:innen solle sich, so Wineburg, folglich an den Strategien professioneller Historiker:innen orientieren. Im britischen Raum entwickelten Lee und Ashby (2000) ein Progressionsmodell zum Aufbau von sogenannten historischen *Second-Order-Concepts*<sup>4</sup>. Verständnis und Anwendung dieser Second-Order-Concepts ist vergleichbar mit dem historischen Denkprozess. Ein Second-Order-Concept ist beispielsweise die differenzierte Verwendung der Konzepte „Quelle“ und „Darstellung“. Die kanadischen Wissenschaftler Peter Seixas und Tom Morton (2013) orientierten sich an der britischen, amerikanischen wie deutschen Diskussionslinie und identifizierten sechs Second-Order-Concepts, wie beispielsweise ein Verständnis für den Umgang mit historischer Evidenz, die für das historische Denken von zentraler Bedeutung sind (siehe auch Seixas, 2017). Jannet van Drie und Carla van Boxtel (2008) prägten die Forschungsdiskussion durch das Rahmenmodell *historical reasoning* von Schüler:innen. Sie unterschieden sechs Kompetenzen von Schüler:innen, die das historische Argumentieren ausmachen, beispielsweise Kontextualisierung und die Verwendung von Quellen.

Das *FUER-Kompetenz-Strukturmodell*<sup>5</sup> historischen Denkens ist ein differenziertes Kompetenzmodell aus dem deutschsprachigen Raum, das den Anspruch formuliert, sowohl für die schulische wie die außerschulische Bildung Anwendung zu finden (Schreiber et al., 2006). Es umfasst vier Kompetenzbereiche, die gemeinsam die globale Kompetenz historischen Denkens konstituieren (Schreiber et al., 2006). Grundlage des Kompetenzmodells ist der Regelkreis des Geschichtsbewusstseins nach Hasberg und Körber (2003), die die Überlegungen von Jörn Rüsen und Karl-Ernst Jeismann zum Geschichtsbewusstsein in ein dynamisches Modell überführten (siehe Abbildung 2 im Hintergrund). Dem Modell zufolge beginnt jeder historische Denkprozess, idealtypisch auch eine Geschichtsunterrichtsstunde oder Unterrichtsreihe, mit

---

<sup>4</sup> Lee und Ashby (2000) grenzten die Second-Order-Concepts von den First-Order-Concepts ab, die bis dato die geschichtsdidaktische Diskussion prägten und ein Verständnis für Konzepte wie Herrschaft, Revolution oder Proletariat umfasste. Die Second-Order-Concepts rückten fortan ein Verständnis auf einer abstrakteren Ebene in den Vordergrund und umfassten auch deren Anwendung, z. B. Evidenz oder Ursache (siehe. auch Lee, Ashby & Dickinson, 1996).

<sup>5</sup> Das Akronym steht für **F**örderung **u**nd **E**ntwicklung von reflektiertem und (selbst-)reflexivem Geschichtsbewusstsein, ein internationales Forschungsprojekt, das sich unter anderem mit Grundlagenforschung zu historischen Kompetenzen beschäftigte.

einer Verunsicherung in der Gegenwart, auf die eine historische Frage folgt. Diese kann entweder an die Vergangenheit gerichtet werden und durch Quellenvergleiche und die Entwicklung einer eigenen historischen Narration bearbeitet werden. Oder die Frage kann an die Geschichte, an vorhandene historische Darstellungen gerichtet werden, die kritisch analysiert werden. Der Erkenntnis oder Urteilsbildung über beispielsweise ein historisches Ereignis, eine historische Entscheidung oder einen Zusammenhang in der Geschichte folgt die Orientierung in der Gegenwart und somit eine Überführung der anfänglichen Verunsicherung in Versicherung (Hasberg & Körber, 2003). Das FUER-Modell legt auf diesen Regelkreis vier Kompetenzbereiche, die zusammen die Kompetenz historischen Denkens ausmachen (siehe Zylinderformen in Abbildung 2).

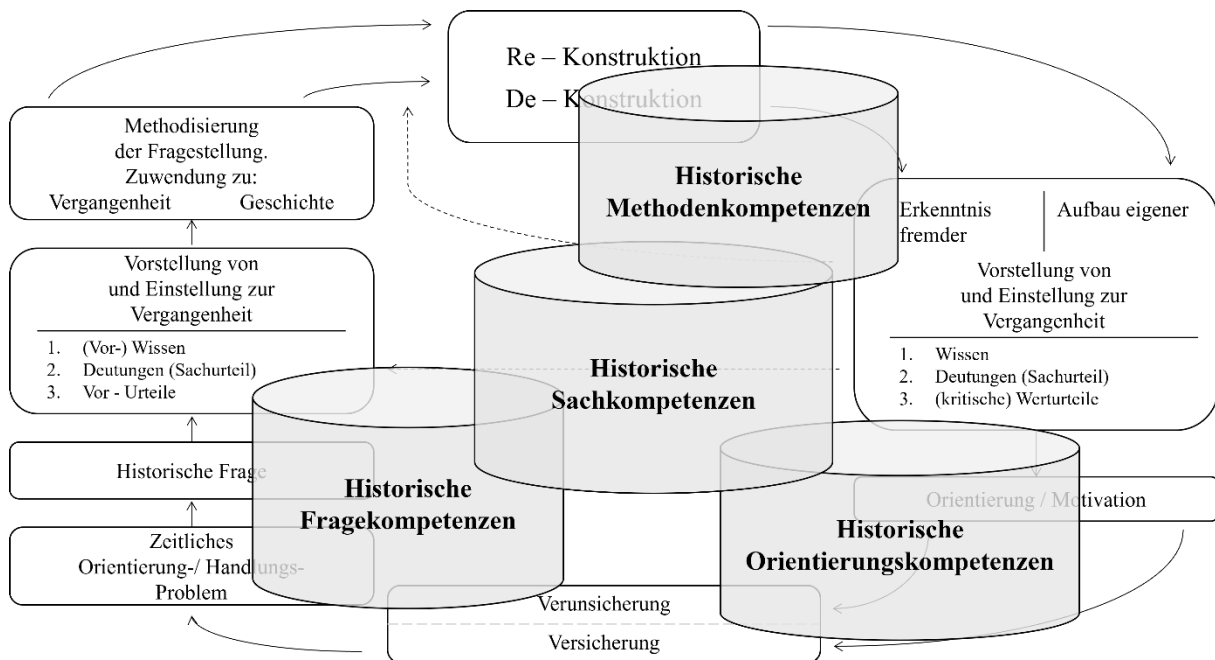


Abbildung 2. Regelkreis historischen Denkens nach Hasberg und Körber (2003) im Hintergrund sowie FUER-Kompetenz-Strukturmodell historischen Denkens (Schreiber et al., 2006)<sup>6</sup>

Die historischen *Fragekompetenzen* beinhalten die Fähigkeiten, eigene historische Fragen zu formulieren und/oder in vorhandenen Darstellungen zu erschließen. Die historischen *Methodenkompetenzen* umfassen die Operationen der Re- und Dekonstruktion: Im Rahmen eines Rekonstruktionsprozesses werden verschiedene Quellen aus der Vergangenheit verglichen, interpretiert und in einen Zusammenhang gesetzt, und so wird eine eigene Darstellung von Geschichte geschaffen. Die Rekonstruktion ist folglich ein „Akt der Synthese“ (Schreiber et al., 2006, S. 24), während es sich beim Dekonstruieren um einen „analytischen Akt“

<sup>6</sup> Die Abbildung ist eine eigene Nachbildung des FUER-Kompetenz-Strukturmodells, um den Regelkreis historischen Denkens besser lesen zu können.

(Schreiber et al., 2006, S. 24) handelt. Hier werden historische Darstellungen entlang einer methodisierten Vorgehensweise kritisch analysiert. Der Kompetenzbereich der historischen *Orientierungskompetenzen* umfasst die Fähigkeiten, die Erkenntnisse aus dem historischen Denkprozess auf die eigene Person, Weltsicht und Lebenswelt zu beziehen. Im Bereich der historischen *Sachkompetenzen*, der anders als die anderen Kompetenzbereiche nicht in Prozesse, sondern durch Konzepte und Kategorien strukturiert ist, schlagen sich die Fähigkeiten nieder, fachspezifische Begriffe und Konzepte differenziert zu verwenden (Schreiber et al., 2006). Hierunter fällt beispielsweise die Unterscheidung von Vergangenheit und Geschichte: Während die Vergangenheit die Summe der Ereignisse in der vergangenen Zeit bezeichnet, ist Geschichte die Rekonstruktion der Vergangenheit und kann lediglich ein (mehr oder weniger gesichertes) Abbild der Vergangenheit sein (z. B. Trautwein et al., 2017).

### ***Qualitätvoller Geschichtsunterricht***

Die angeführten Ziele und Kompetenzmodelle zeigen auf, welche Erträge der Geschichtsunterricht bringen soll. Damit der Unterricht das historische Denken der Schüler:innen fördert, muss er qualitativ gestaltet sein. Zur Beschreibung des Unterrichts hat sich in der deutschsprachigen Unterrichtsforschung eine Unterscheidung zwischen Oberflächen- und Tiefenmerkmalen etabliert (z. B. Decristan, Hess, Holzberger & Praetorius, 2020). Die Oberflächenmerkmale umfassen jene Gestaltungsmerkmale des Unterrichts, die gut beobachtbar sind, wie zum Beispiel Methoden, Sozialformen oder Medien (Kunter & Trautwein, 2013). Tiefenmerkmale umfassen die tatsächlichen Lehr-Lern- und Interaktionsprozesse, die von außen nicht direkt erkennbar sind. Diese Lehr-Lernprozesse wurden in der pädagogisch-psychologischen Unterrichtsforschung lange mithilfe des Rahmenmodells der drei Basisdimensionen (Klieme, Schümer & Knoll, 2001) verstanden und erforscht, welches die drei Merkmale (1) Klassenführung, (2) konstruktive Unterstützung durch die Lehrkraft und (3) das Potenzial zur kognitiven Aktivierung unterscheidet. Klassenführung beschreibt, wie gut der Unterricht von der Lehrkraft gesteuert ist und ob Ziele klar formuliert werden, damit alle Schüler:innen möglichst viel Lernzeit effektiv nutzen können. Das Tiefenmerkmal der konstruktiven Unterstützung umfasst methodisch-didaktische sowie sozio-emotionale Anteile von Unterstützung (Trautwein et al., 2021) und beschreibt den wertschätzenden Umgang zwischen Lernenden und Lehrkraft. Das Potenzial zur kognitiven Aktivierung beschreibt die Tiefe der Lernprozesse durch anregende Aufgaben und Unterrichtssituationen wie beispielsweise kognitive Konflikte. Empirische Studien konnten diese dreifaktorielle Struktur vielfach bestätigen (z. B. Kunter et al., 2008; Wagner, Göllner, Helmke, Trautwein & Lüdtke, 2013; Wagner et al.,

2016) und den Tiefenmerkmalen im Vergleich zu den Oberflächenmerkmalen eine höhere Erklärungskraft für den Lernerfolg nachweisen (z. B. Hattie, 2009; Kunter & Voss, 2011; Seidel & Shavelson, 2007).

Kürzlich erweiterten Praetorius et al. (2020) das Modell der Basisdimensionen auf insgesamt sieben Dimensionen in einem Syntheseframework, das die Merkmale der Unterrichtsqualität differenzierter aufschlüsselt und eine Unterscheidung zwischen Angebots- und Nutzungsaspekten vornimmt. Bisher stützt sich die empirische Unterrichtsqualitätsforschung vor allem auf Daten aus dem Mathematikunterricht. Zur Verknüpfung der generischen Tiefenmerkmale mit fachspezifischen Unterrichtsmerkmalen des Geschichtsunterrichts gibt es zum jetzigen Zeitpunkt wenige, vor allem theoretische und konzeptuelle Überlegungen (Trautwein et al., 2021; Zülzdorf-Kersting, 2020), obwohl die interdisziplinären Projekte für Erkenntnisprozesse in beiden Disziplinen fruchtbar erscheinen (Trautwein et al., 2021).

Dabei gibt es in der Geschichtsdidaktik etwa seit 2006 vereinzelt Bemühungen, Qualitätsmerkmale guten Geschichtsunterrichts theoretisch und empirisch zu fundieren (Barricelli & Sauer, 2006; Gautschi, 2009; Meyer-Hamme, Thünemann & Zülzdorf-Kersting, 2016; Zülzdorf-Kersting, 2020). Gautschi (2009) führte eine umfangreiche Unterrichtsstudie durch und kam zu dem Schluss, dass guter Geschichtsunterricht nicht ausreichend von fachunspezifischen Qualitätsmerkmalen beurteilt werden kann. Vielmehr müssten immer auch fachspezifische Komponenten Berücksichtigung finden. Geschichtsunterricht ist nach Gautschi dann *gut*, wenn zwei Schlüsselfaktoren erfüllt sind: ein Bezug zur Lebenswelt der Schüler:innen (Lerngegenstand: Inhalte, Themen, Medien) sowie aktivierende Lerngelegenheiten. Gautschi (2009) schlussfolgerte weiter, dass für den Lerngegenstand und die Lernaktivität eher fachspezifische Gütekriterien angewendet werden, während fachunspezifische Gütekriterien, wie die drei Basisdimensionen, für die Unterrichtsgestaltung durch die Lehrkraft gelten sollten.

Zülzdorf-Kersting (2020) erarbeitete einen Überblick, wie Unterrichtsqualität in geschichtsdidaktischen Studien bisher thematisiert wurde und wie diese Merkmale mit den generischen Merkmalen des Syntheseframeworks nach Praetorius et al. (2020) verknüpft werden können. Auf der Grundlage empirischer Studien zum Geschichtsunterricht arbeitete er neun Merkmale von qualitativem Geschichtsunterricht heraus, von denen sich acht Merkmale in das Syntheseframework, und zwar in die Dimensionen der Themenauswahl (z. B. Relevanz des Themas) und kognitiven Aktivierung (z. B. gute Lernaufgaben) einordnen ließen. Für das neunte Merkmal, eine Förderung des Unterrichtsdiskurses, schlug er eine querschnittliche Erweiterung des Syntheseframeworks vor. Zülzdorf-Kersting (2020) zeigte zudem auf, dass sich

die geschichtsdidaktischen Studien bisher nicht mit den Merkmalen der Klassenführung und konstruktiven Unterstützung als Tiefenmerkmale des Unterrichts beschäftigt.

Trautwein et al. (2021) verknüpften in einem Positionspapier jüngst die drei Basisdimensionen von Unterrichtsqualität mit dem FUER-Kompetenz-Strukturmodell und stellten fest, dass die drei Basisdimensionen, die teilweise auch bereits im FUER-Modell angelegt sind, wertvolle Anknüpfungspunkte für die Beurteilung der Unterrichtsqualität im Fachunterricht darstellen. Die Autor:innen erkannten aufgrund der interdisziplinären Perspektive einen Erkenntnisgewinn für beide Disziplinen: Die Basisdimensionen zeigen der Geschichtsdidaktik auf, anhand welcher Merkmale die Qualität des Unterrichts bestimmt werden kann. Die Besonderheiten des Geschichtsunterrichts zeigen der Unterrichtsforschung, dass qualitativvoller Unterricht insbesondere durch die Verbindung zweier Dimensionen erreicht werden kann. Trautwein et al. schrieben dazu: „Erkenntnisgewinn im Geschichtsunterricht kann Schüler:innen dagegen [Anm. d. V.: im Vergleich zu Mathematik] in Mark und Bein gehen. Man kann sich deshalb guten Geschichtsunterricht, der zu wirklicher Orientierung führt, nur bei einer gelungenen Synthese von kognitiver Aktivierung und sozio-emotionaler konstruktiver Unterstützung denken“ (S. 269).

Zusammenfassend zeigt sich, dass in westlichen Demokratien das Ziel des Geschichtsunterrichts darin besteht, Schüler:innen zu historischem Denken zu befähigen und sie somit auch zu mündigen Bürger:innen der Gesellschaft auszubilden. Dies gelingt mit der Gestaltung von qualitativvollem Geschichtsunterricht, wie insbesondere theoretische und konzeptuelle Überlegungen durch die Verknüpfung von Geschichtsdidaktik und Unterrichtsforschung zeigen. Dabei wird deutlich, dass der historische Denkprozess mit dem Ziel, *wirkliche Orientierung* in der Gegenwart zu finden, und ein entsprechender Geschichtsunterricht, der Schüler:innen den Raum für Identitätsarbeit (Bracke & Flaving, 2018) gibt, mehr ist als ein rein kognitiver Akt: „Emotionen [...] formen Bewusstsein, Wahrnehmung und Erinnerung und sind damit eine zentrale Kategorie des historischen Lernens, der historischen Sinnbildung“ (Brauer & Lücke, 2013, S. 19). Emotionen sind aber nicht nur wesentlicher Teil der Wahrnehmung und des Lernprozesses. Sie spielen im Unterrichtsgeschehen auf verschiedenen Ebenen eine zentrale Rolle und können im Angebot-Nutzungs-Modell an unterschiedlichen Stellen verortet werden. Im folgenden Abschnitt werde ich zunächst kurz auf die Erforschung von Emotionen im Unterricht eingehen, ehe ich die spezifische Rolle der Emotionen beim historischen Lernen thematisiere.



### 1.1.3 Emotionen im (Geschichts-)Unterricht

Emotionen spielen eine zentrale Rolle im Unterrichtsgeschehen (Pekrun & Linnenbrink-Garcia, 2014) und lassen sich an drei Schlüsselstellen im Angebot-Nutzungs-Modell verorten (Hascher & Brandenberger, 2018): in den Lernvoraussetzungen der Schüler:innen (z. B. Motivation, Interesse), in den Wirkungen des Unterrichts (z. B. Freude über eine gute Note, Motivation) und in den Lernaktivitäten, in denen Emotionen wie beispielsweise Neugier oder Langeweile den Lernprozess begleiten. Grundsätzlich lassen sich Emotionen anhand von fünf Komponenten beschreiben: einer affektiven, kognitiven, physiologischen, expressiven und motivationalen Komponente (Goetz, Zirngibl, Pekrun & Hall, 2003; Hascher & Brandenberger, 2018; Pekrun, 2006). Hinsichtlich des Unterrichtskontextes differenziert die pädagogisch-psychologische Forschung Emotionen zudem danach, auf welchen Gegenstand sich die Emotionen beziehen. Unterschieden werden Leistungsemotionen, epistemische Emotionen, gegenstandsbezogene Emotionen und soziale Emotionen (Loderer, Pekrun & Plass, 2020; Pekrun, 2016). Während Leistungsemotionen als Wirkung des Unterrichts konzeptualisiert werden können, begleiten epistemische (den Lernprozess betreffende) und gegenstandsbezogene (den Lerninhalt betreffende) Emotionen die Lernaktivitäten. Sie sind auf kognitive Dissonanzen und den Unterrichtsinhalt gerichtet. Soziale Emotionen finden in Interaktionsprozessen statt und sind deshalb nicht an einer konkreten Stelle im Angebots-Nutzungs-Modell zu verorten. Speziell im Geschichtsunterricht finden sich Emotionen zudem an einer weiteren Stelle: Sie sind als Teil des unterrichtlichen Angebots auch Gegenstand des historischen Lernens (Brauer, 2019; Brauer & Lücke, 2022; Hasberg, 2013). Schüler:innen lernen über die Gefühle historischer Akteur:innen, Gefühlskonzepte sowie die Historizität von Gefühlen: Was empfanden die Menschen aus Ost- und Westdeutschland, als die Mauer fiel? Wie wurde Heimatgefühl in der DDR propagiert? Fühlte sich die Fluchterfahrung für die Menschen aus der ehemaligen DDR anders an als für Menschen, die heute fliehen?

Das Ziel des Geschichtsunterrichts, eine Orientierung für Gegenwart und Zukunft zu stiften, zeigt deutlich, dass insbesondere gegenstandsbezogene Emotionen auch auf der Ebene der Lernenden, der Subjektebene (Hasberg, 2013), im Geschichtsunterricht „wesentlich zur Dynamik jedes Erkenntnisprozesses selber“ (Rüsen, 2020, S. 149) gehören. Schüler:innen reagieren möglicherweise emotional auf historische Ereignisse, grenzen sich ab, orientieren sich, bilden ihre Persönlichkeit und Identität aus (z. B. Helmsing, 2014; Meyer-Hamme, 2013). Dass Emotionen der Schüler:innen eine zentrale Rolle im historischen Denkprozess spielen, spiegelt sich in den Forschungsarbeiten der deutschsprachigen Geschichtsdidaktik trotz fortwährender Diskussionen erst seit knapp zehn Jahren wider. Das hat auch historische Gründe: So wurde der

Geschichtsunterricht im Kaiserreich und während der Macht der Nationalsozialisten zielführend genutzt, um Schüler:innen zu indoktrinieren (Brauer, 2019; Frevert, 2011, 2018). Dies führte zu einer stark kognitiv-geprägten Debatte um das historische Lernen, obwohl Gefühle in den Grundlagenwerken Jeismanns (1977, 1994) und Rüsens (1994) doch eine Berücksichtigung fanden. Nachdem auch nach einer geschichtsdidaktischen Tagung am Georg-Eckert-Institut für internationale Schulbuchforschung in Braunschweig 1991 weiterhin nur wenige Forschungsarbeiten zu Emotionen durchgeführt wurden, fand 2011 erneut eine solche Tagung am Max-Planck-Institut für Bildungsforschung mit dem Arbeitsbereich Didaktik der Geschichte der Freien Universität Berlin statt. Auf diese Tagung folgten mittlerweile vielfache Bemühungen, die Rolle von Emotionen im historischen Denkprozess zu ergründen (Ballis & Gloe, 2019; Brauer, 2016; Brauer & Lücke, 2013, 2022; Brüning, 2018; Frevert, 2018). In ihrer *Theorie des Geschichtsunterrichts* schlugen Bracke et al. (2018) die explizite Anerkennung von Emotionen im historischen Lernprozess vor und begegneten damit hinsichtlich ihrer Berücksichtigung einem Mangel an bisherigen Modellen historischen Denkens, so die Autor:innen. Auch international wurden Emotionen als Teil des Geschichtsunterrichts und des historischen Lernens berücksichtigt (z. B. Sheppard, Katz & Grosland, 2015): Lernen über Geschichte heißt auch, sich mit Krieg und Gewalt in der Vergangenheit zu beschäftigen (z. B. Levy & Sheppard, 2018); es heißt, die Relevanz der Vergangenheit für das persönliche Leben zu erkennen (z. B. Van Straaten et al., 2016; Van Straaten, Wilschut, Oostdam & Fukkink, 2019) oder Empathie für die historischen Akteur:innen zu empfinden (z. B. Baron, 2012; Endacott & Brooks, 2013).

Allerdings lassen sich in der Praxis Objekt- (Gefühle in der Geschichte als Lerngegenstand) und Subjektebene (Gefühle der Lernenden) der Gefühle nur schwer voneinander trennen. Das wird beispielsweise am Konstrukt *Historical Empathy* (Brooks, 2009; Endacott, 2010; Endacott & Brooks, 2013) deutlich: Lernende sollen die Perspektive historischer Akteur:innen verstehen, sie in den historischen Kontext einordnen und eine Verbindung zu eigenen, wenn auch unterschiedlichen Erfahrungen herstellen (Endacott & Brooks, 2013). Endacott und Brooks (2018) wiesen in ihren konzeptuellen Überlegungen bereits darauf hin, dass die emotionale Reaktion oft mit einer Übernahme des emotionalen Zustands der historischen Akteur:innen einhergehe und Regulationsstrategien erfordere. Brauer (2019) bezeichnete eine solche Verzweigung von Emotionen auf Subjekt- und Objektebene als Emotionalisierung oder gar emotionale Überwältigung. Ein Nachempfinden und gänzlich Verstehen der Perspektiven könne nicht gelingen, da wir heute einen anderen „Erfahrungsraum und Erwartungshorizont“ (Koselleck, 1979, S. 349, zitiert nach Brauer, 2019) haben als die historischen Akteur:innen damals. Werden Emotionen von den Lernenden sehr intensiv erlebt, müssen die Lernenden

Anstrengungen dafür aufwenden, diese zu regulieren (Bracke & Flaving, 2018; Endacott & Brooks, 2018). Das könne möglicherweise dazu führen, dass weniger kognitive Kapazitäten für das eigentliche Lernen zur Verfügung stehen und somit historisches Lernen blockiert wird bzw. die Schüler:innen überwältigt werden. Speziell in Deutschland dürfe diese emotionale Überwältigung im Rahmen schulischer und außerschulischer historisch-politischer Bildung allerdings nicht stattfinden: Im Beutelsbacher Konsens einigten sich Akteur:innen aus der politischen Bildung auf ein sogenanntes Überwältigungsverbot, das bis heute einen Grundsatz der historisch-politischen Bildung darstellt (Krüger, 2020; Wehling, 1977).

Obwohl es sowohl in der internationalen wie deutschsprachigen Forschungsgemeinschaft mittlerweile *Common Sense* zu sein scheint, dass Emotionen Teil des historischen Denkprozesses sind, bleiben die Fragen offen, inwiefern Emotionen in Bezug zur Förderung historischen Denkens stehen (Bracke & Flaving, 2018; Brauer & Lücke, 2022; Endacott, 2010; Kohlmeier, 2006; Meyer-Hamme, 2013) und wann Emotionen den Zugang zur Geschichte verschließen (Brauer & Lücke, 2013). Ein Lernarrangement im Geschichtsunterricht macht diese ambivalente Rolle von Emotionen im historischen Denkprozess besonders deutlich: die Begegnung mit Zeitzeug:innenberichten. Die persönlichen Geschichten aus der Vergangenheit erzählen häufig von Gefühlen, fordern Schüler:innen zum Nacherleben auf, begeistern für das Fach und beinhalten eine Botschaft für das Leben in Gegenwart und Zukunft. Gleichzeitig genießen Zeitzeug:innenberichte ein größeres Vertrauen, *wahre* Geschichten über die Vergangenheit erzählen zu können (Angvik & von Borries, 1997), was eine Fehlvorstellung von Geschichte darstellt. Im folgenden Kapitel werde ich das Lernen mit Zeitzeug:innenberichten als Unterrichtsangebot thematisieren und empirische Forschung zu diesem Lernarrangement vorstellen.

## **1.2 Unterrichtsangebot: Das Lernarrangement mit Zeitzeug:innenberichten**

Im folgenden Kapitel werde ich zunächst konzeptionelle Überlegungen zur Zeitzeug:innenschaft anstellen, die Grundlage der vier Dissertationsstudien sind. Daran anschließend beschreibe ich das Lernarrangement mit Zeitzeug:innenberichten aus der Perspektive der Empirischen Bildungsforschung und stelle bisherige empirische Studien zu dem Lernarrangement vor. Abschließend skizziere ich anhand der empirischen Studien relevante Oberflächen- und Tiefenmerkmale des Unterrichts mit Zeitzeug:innenberichten und weise somit Gelingensfaktoren für das Lernarrangement aus.

### **1.2.1 Zeitzeug:innenberichte als Lernmaterial: konzeptuelle Überlegungen**

Albrecht Möller und Mike Michelus haben beide in der ehemaligen DDR gelebt – und waren dort inhaftiert. Albrecht Möller wurde aufgrund eines gescheiterten Fluchtversuchs aus der ehemaligen DDR festgenommen und für vier Jahre von seiner Familie getrennt, als seine Tochter vier und sein Sohn zwei Jahre alt waren. Mike Michelus war Teil der Punkszene in der ehemaligen DDR und wurde während der Teilnahme an einem Kunstprojekt auf dem Berliner Alexanderplatz festgenommen und zu einem Jahr Haft verurteilt. Während Möller selten vor Schüler:innen über seine Erfahrungen spricht, ist Michelus Hobby-Schauspieler, hat seinen Vortrag mit Musik, Objekten und Bildern ausgeschmückt und bereits hunderte Male vorgetragen. Beide sind als Zeitzeugen in der Datenbank [www.ddr-zeitzeuge.de](http://www.ddr-zeitzeuge.de) registriert – ein Portal, das Zeitzeug:innen mit Hafterfahrung in der ehemaligen DDR an Schulen vermittelt und ihren Besuch finanziert. Sie werden mit dem Ziel in den Unterricht eingeladen, „DDR-Geschichte hautnah“ zu erleben und „eine fremde, unbegreifliche Welt“ für die Jugendlichen verständlich und lebendig zu machen ([www.ddr-zeitzeuge.de](http://www.ddr-zeitzeuge.de)). Die Lebensgeschichten von Möller und Michelus und ihre Art, diese in der Gegenwart zu erzählen, verdeutlichen die Individualität der persönlichen Geschichten und Charaktere. Zwar suggeriert das Thema der Zeitzeug:innenbörse (Menschen mit Hafterfahrung in der ehemaligen DDR) eine Art Vergleichbarkeit zwischen den Berichten, die konkreten Geschichten unterscheiden sich allerdings auf vielen Ebenen: Unterschiedlich sind beispielsweise der Grund für die Inhaftierung, die Lebensphase, in der die beiden Männer zum Zeitpunkt der Inhaftierung waren, ihr Stand der persönlichen Aufarbeitung der Erfahrung sowie die Sicherheit, mit der sie von ihren Erfahrungen berichten. Es wird offensichtlich, dass Möller und Michelus nicht die gleiche Perspektive auf die Vergangenheit haben und Lernenden im Geschichtsunterricht folglich nur ihren individuellen Blick bieten können.

Es unterscheiden sich aber nicht nur die Lebensgeschichten selbst, sondern auch, welche Funktion Zeitzeug:innen<sup>7</sup> in verschiedenen Bereichen der Geschichtskultur innehaben. In der Geschichtswissenschaft folgt die *Oral-History*-Methode der Frage, wie die „kleinen Leute“ mit Herausforderungen ihrer Zeit umgingen. Ziel dieses Zugangs zur Vergangenheit ist es, historische Narrative nicht ausschließlich aus der Perspektive der Mächtigen zu entwickeln, sondern den Alltagserfahrungen und Lebensgeschichten der „kleinen Leute“ eine Stimme zu geben (Bertram, 2017). Zeitzeug:innen erzählen in dieser Tradition häufig in lebensgeschichtlichen Interviews von ihren persönlichen Erfahrungen. Diese werden aufgezeichnet und kriteriengeleitet ausgewertet (Wierling, 1997). Im Gegensatz dazu haben Zeitzeug:innen in der Erinnerungskultur, wie beispielsweise in Fernsehdokumentationen, Gedenkstätten oder der außerschulischen Bildungsarbeit eine andere Funktion. Zum Beispiel werden in der *Holocaust Education*, der historischen und politischen Bildung zur Shoah, Zeitzeug:innenberichte – häufig in Form von Opfererzählungen über nationalsozialistische Verbrechen – genutzt, um nachfolgende Generationen zu sensibilisieren und vor einer Wiederholung der Schreckensgeschichte der Shoah zu warnen (Bertram, 2017; Brüning, 2018). Auch unabhängig vom Bezug zum Holocaust werden Zeitzeug:innenberichte in der Erinnerungskultur häufig mit dem Ziel angeführt, ein bestimmtes historisches Narrativ, d. h. eine bestimmte Darstellung der Vergangenheit, zu autorisieren (Bertram, 2017; Sabrow, 2012). Im Geschichtsunterricht mit Zeitzeug:innenberichten finden sich hingegen alle der hier aufgeführten Funktionen wieder (Bertram, 2017): Zeitzeug:innen sollen sensibilisieren, Interviewpartner:innen für Oral-History-Projekte sein oder, wie Möller und Michelus, die Vergangenheit durch die individuellen Geschichten lebendig machen.

In Anbetracht dieser Unterschiedlichkeit von Zeitzeug:innenberichten auf den verschiedenen Ebenen erscheint es fast unmöglich, allgemeine Aussagen über Zeitzeug:innenberichte treffen zu können. Allerdings finden sich in der Literatur Bemühungen zu einem generischen Verständnis von Zeitzeug:innenschaft (Bertram, 2020; De Jong, 2018; Ernst, 2012). In Anlehnung an die Unterscheidung zwischen Vergangenheit (die vergangenen Ereignisse) und Geschichte (das Abbild der Vergangenheit durch Rekonstruktionen; siehe Kapitel 1.1.2) wird zwischen Zeitzeug:innen der Vergangenheit (*witness of the past*; implizite Zeitzeug:innen) und Zeitzeug:innen der Geschichte (*witness to history*; explizite Zeitzeug:innen) unterschieden (De

---

<sup>7</sup> Die Ursprünge des Konzepts der Zeitzeug:innenschaft in Deutschland werden von Historiker:innen auf den Eichmann-Prozess 1961 und die Auschwitz-Prozesse 1963 bis 1968 zurückgeführt, da hier Menschen öffentlich die Verbrechen der Nationalsozialisten vor Gericht bezeugen sollten (vgl. Bertram, 2017; Brüning, 2018; Ernst, 2012). Explizit verwendet wurde der Begriff des Zeitzeugen allerdings erst in einer Buchbesprechung des Spiegel 1975 (vgl. Bertram, 2017; Brüning, 2018).

Jong, 2018; Ernst, 2012). Zeitzeug:innen der Vergangenheit sind all jene Menschen, die ein historisches Ereignis erfahren, gesehen oder gehört haben. Diese werden zu Zeitzeug:innen der Geschichte, wenn sie ein Zeugnis über ihre Erfahrungen ablegen und damit eine Darstellung vornehmen (De Jong, 2018; Ernst, 2012). De Jong (2018) erarbeitete fünf idealtypische Charakteristika, die Zeitzeug:innen der Geschichte auszeichnen und für die sie sich an der Person der juristischen Zeug:in orientierte: (1) Ziel des Zeugnisses ist Bildung, neue Informationen über die Vergangenheit zu vermitteln oder emotional zu berühren; (2) die Zeugnisse zeichnen sich durch räumliche und zeitliche Präsenz der Person aus; (3) die subjektiven Erfahrungen werden in Worte gefasst, interpretiert und (4) einem Publikum präsentiert – das kann unmittelbar sein in einem Vortrag oder auch über einen videografierten oder verschriftlichen Bericht, der mittelbar an ein Publikum gelangt; (5) Zeitzeug:innen benötigen die Zuschreibung der Glaubwürdigkeit.

Auch Bertram (2020) arbeitete drei grundsätzliche Merkmale von Zeitzeug:innenberichten heraus, welche die „Komplexität des historischen Denkprozesses“ (S. 252) verdeutlichen können: Erinnerung, Narrativität und Authentizität. (1) Das Merkmal der Erinnerung beschreibt, dass jede Erzählung von Zeitzeug:innen eine subjektive Interpretation der Erfahrung darstellt, die sich mit zunehmender Distanz zum Ereignis verändern kann. (2) Das Merkmal der Narrativität beschreibt die Verbindung entlang der Zeitachse, denn ein Zeitzeug:innenbericht ist immer eine Erzählung aus der Gegenwart über die Vergangenheit und hat in der Regel eine (implizite oder explizite) Botschaft für die Zukunft. (3) Das Merkmal der (vermeintlichen) Authentizität des Berichts beschreibt, dass Zeitzeug:innen von etwas berichten, das sie selbst gehört, gesehen oder erlebt haben (Bertram, 2020). Die Konzeptualisierungen von De Jong (2018) und Bertram (2020) lassen sich zusammenführen: Das Merkmal der Authentizität lässt sich in De Jongs (2018) Beschreibungen bei räumlicher und zeitlicher Präsenz sowie der Zuschreibung von Glaubwürdigkeit einordnen. Das Merkmal der Erinnerung erweitert die Eigenschaft der subjektiven Erinnerung nach De Jong (2018) um die Dynamik, dass sich Interpretationen der Erfahrungen über die Zeit verändern. Das Merkmal der Narrativität erweitert das Merkmal der zeitlichen und räumlichen Präsenz nach De Jong (2018), in dem es die Verknüpfung zur Gegenwart und Zukunft herstellt.

In Anlehnung an Bertram (2020) und De Jong (2018) verstehe ich in der vorliegenden Dissertation Zeitzeug:innenberichte (der Geschichte) als individuelle Erfahrungsberichte über die Vergangenheit, die einem Publikum mit einem bestimmten Ziel zugänglich gemacht werden und denen aufgrund räumlicher und zeitlicher Präsenz eine besondere Glaubwürdigkeit attestiert wird. All diese Eigenschaften zusammengenommen, resultieren in einer sogenannten Aura

der Zeitzeug:innen. Diese Aura von Zeitzeug:innen macht nach Sabrow (2007) die didaktische Qualität der Geschichten aus. Ich unternehme in der vorliegenden Dissertation den Versuch, Erkenntnisse über das Lernen mit Zeitzeug:innenberichten allgemein zu gewinnen, obwohl die hier aufgeführte Unterschiedlichkeit der Zeitzeug:innenberichte auch eine Rolle für die Wahrnehmung der Lernenden und das Lernen selbst spielt. Einige der zentralen Dimensionen, anhand derer sich Zeitzeug:innenberichte – insbesondere in ihrem Einsatz im Geschichtsunterricht – unterscheiden, werden im folgenden Abschnitt etwas detaillierter ausgeführt: die didaktischen Zielsetzungen und Themen, Einsatzmöglichkeiten und Formate von Zeitzeug:innenberichten im Unterricht. Im Anschluss stelle ich bisherige empirische Studien zum Lernarrangement vor und verknüpfe abschließend das Lernarrangement mit Zeitzeug:innenberichten mit dem Modell der Basisdimensionen von Unterrichtsqualität aus der Empirischen Bildungsforschung, um Gelingensfaktoren für das Lernarrangement zu identifizieren.

## **1.2.2 Zeitzeug:innenberichte und ihre Einbettung im Unterricht**

„Biografische Zugänge zur Geschichte sind en vogue. Historischen Figuren so nahe wie möglich zu kommen, erscheint bisweilen sogar als ein Königsweg zur attraktiven Darstellung und Vermittlung geschichtlicher Inhalte und für den Erwerb entsprechender Einsichten“ (Memminger, 2022, S. 1). Für das Lernen über und von der Geschichte können Zeitzeug:innenberichte mit ihren individuellen Lebensgeschichten die Distanz der Schüler:innen zur Vergangenheit überbrücken. Wie genau das Lernarrangement mit den biografischen Zugängen zur Geschichte allerdings gestaltet ist, ist beispielsweise abhängig von der didaktischen Zielsetzung und dem Thema, dem Format und der konkreten Einbettung in den Unterricht.

### ***Didaktische Zielsetzungen & Thema***

Zeitzeug:innenberichte können mit unterschiedlichen didaktischen Zielsetzungen in den Unterricht integriert werden und damit unterschiedliche Ebenen des Lernens fördern. Ein zentrales Ziel des Einsatzes von Zeitzeug:innenberichten ist, die von Jugendlichen häufig mit trockenen Zahlen und Fakten in Verbindung gebrachte Geschichte durch die persönlichen Erfahrungsberichte lebendig werden zu lassen und somit die Distanz zur Vergangenheit abzubauen (Ballis & Schwendemann, 2021; Brauer & Lücke, 2022; Memminger, 2022). Der biografische Zugang zur Vergangenheit ermöglicht Lernenden, eine Verbindung zum eigenen Leben herzustellen, und kann so die Relevanz von Geschichte für Gegenwart und Zukunft aufzeigen. Das fördert die Motivation der Lernenden, sich mit historischen Themen auseinanderzusetzen (Bertram et al., 2017; Van Straaten et al., 2016).

Darüber hinaus werden Zeitzeug:innenberichte im Unterricht auch mit dem Ziel eingesetzt, für in der Geschichte begangenes Unrecht oder für Werte wie beispielsweise Menschenrechte und Toleranz zu sensibilisieren (Brüning, 2018; Obens & Geißler-Jagodzinski, 2010). Die Geschichten über Leid und Unrecht regen Lernende dazu an, mit den Zeitzeug:innen mitzufühlen und so ihre Gefühle nachzuempfinden und ihre Gedanken und Entscheidungen nachzuvollziehen. Problematisch ist dabei allerdings, dass diese Berichte ohne angemessene Kontextualisierung zu Betroffenheit der Lernenden führen können, was jedoch kein Lernziel an sich darstellen darf (siehe Kapitel 1.1.3; Brüning, 2018; Heyl, 1995; Wehling, 1977). Zudem argumentierte Brauer (2019), dass der Anspruch des Sich-hinein-Versetzens in vergangene Biografien nicht gelingen kann, da der Erfahrungsraum und die Lebenssituation in der Vergangenheit anders waren, als sie für Lernende heute sind.

Ein weiteres didaktisches Ziel orientiert sich an der Oral-History-Methode aus der Geschichtswissenschaft: Schüler:innen sollen in Projekten den Prozess des historischen Erkenntnisgewinns durchlaufen, nach einer historischen Fragestellung eigenständig Interviews mit Zeitzeug:innen führen, diese Interviews auswerten und präsentieren und die Projektarbeit reflektieren (z. B. Ardemendo & Kuszmar, 2013; Henke-Bockschatz, 2014; Whitman, 2004). Zahlreiche Handbücher und Lehrerhandreichungen unterstützen Lehrkräfte in der Planung dieser „Hochform des Lernens im Geschichtsunterricht“ (Erbar, 2012, S. 18), stellen Potenziale, aber auch Risiken dieser Projektarbeit vor (z. B. Ernst, 2012; Henke-Bockschatz, 2007; Llewellyn & Ng-A-Fook, 2017; Schreiber, 2009; Sochatzy, 2010; Trškan, 2016; Wolter, 2013). Der Unterricht mit Zeitzeug:innenberichten kann somit den historischen Denkprozess anstoßen sowie Einblicke in zentrale historische Prinzipien wie beispielsweise die Multiperspektivität (Berücksichtigung verschiedener Perspektiven) vermitteln – folglich Kompetenzen historischen Denkens fördern (Bertram, 2017; von Plato, 2001).

Je nach Zielsetzung des Einsatzes von Zeitzeug:innenberichten im Unterricht werden Schüler:innen primär kognitiv (Oral History, historischer Denkprozess), motivational (Geschichte wird lebendig) oder emotional (Sensibilisierung) angesprochen. Thematisch eignen sich jedoch nicht alle Zeitzeug:innenberichte für jede der didaktischen Zielsetzungen gleichermaßen. Während in der Holocaust Education die Erfahrungsberichte von Überlebenden des Holocaust und weiteren Opfern nationalsozialistischer Verbrechen dazu dienen können, Schüler:innen für das Unrecht in der Vergangenheit zu sensibilisieren und den anonymen Zahlen „ein Gesicht zu geben“, eignen sich diese Berichte nicht dazu, sie im Sinne des historischen Denkprozesses kritisch zu analysieren, insbesondere dann nicht, wenn die Person den Unterricht persönlich besucht (Herzing & Longerich, 2013). Anders verhält es sich mit Menschen,



die von ihren Alltagserfahrungen aus einer bestimmten Zeit berichten. Hier können im Sinne der Oral-History-Methode verschiedene Perspektiven verglichen, hinterfragt und somit ein Lernarrangement gestaltet werden, in dem Kompetenzen historischen Denkens gefördert werden können (Bertram, 2017). Allen Zeitzeug:innenberichten ist allerdings gemein, dass sie den Schüler:innen durch die persönlichen Biografien einen individuellen Zugang zur vergangenen Zeit eröffnen und somit abstrakte Daten und Ereignisse mit konkreten Lebensgeschichten füllen. Für das Lernen über Geschichte ist dabei die Erkenntnis zentral, dass Zeitzeug:innenberichte nur einzelne Ausschnitte der Geschichte sind und nicht das große Ganze darstellen (Brauer, 2017).

### **Formate**

Der Einsatz von Zeitzeug:innenberichten im Unterricht unterscheidet sich aber nicht nur mit Blick auf das didaktische Ziel und das Thema, sondern auch im Format und der didaktischen Einbettung der Berichte in den Unterricht. Verschriftlichte Berichte von Zeitzeug:innen in Schulbüchern stellen in der Regel eine Textart von vielen Textgattungen dar, mit dem Schüler:innen in der Gesamtschau des Schulbuchkapitels etwas über die Vergangenheit lernen sollen. Diese Form der Zeitzeug:innenberichte sticht wenig aus dem üblichen Unterrichtsgeschehen heraus (Brüning, 2018; Hogervorst, 2018), obwohl auch hier der biografische Zugang zur Vergangenheit im Vordergrund steht. Darüber hinaus werden Zeitzeug:innenberichte auch in Form von Videoclips in den Unterricht integriert. Ausgehend von den audiovisuellen Aufzeichnungen Holocaust-Überlebender entstand mit der Gründung des *Fortunoff Video Archive for Holocaust Testimonies* ab 1979 eine Art „Digital Turn“ in der Zeitzeug:innenarbeit (Brüning, 2018), der nicht nur für die Oral-History-Methode, sondern auch für die pädagogische Arbeit fortan neue Möglichkeiten darstellte. In den vergangenen Jahren ist die Zahl der Online-Archive von videografierten Zeitzeug:innenberichten stark gewachsen. Das internationale *Visual History Archive* der *USC Shoah Foundation* ist mit 55 000 Interviews von Überlebenden des Holocaust das größte seiner Art (<https://vhaonline.usc.edu/>). Das *Zeitzeugenportal* umfasst Videoclips von 1 000 Interviews mit Zeitzeug:innen vom Ersten Weltkrieg bis zur Gegenwart (<https://www.zeitzeugen-portal.de/>). Das *Zeitzeug:innenportal* der *Beaufragten des Landes Brandenburgs zur Aufarbeitung der kommunistischen Diktatur* stellt Zeitzeug:inneninterviews zusammen, die über politisches Unrecht und Widerstand von 1945 bis 1989 erzählen (<http://www.zeitzeugen.brandenburg.de/>). Diese drei Beispiele zeigen die umfassende Verfügbarkeit videografierten Erfahrungsberichte, derer sich Lehrkräfte für ihren Un-

terricht bedienen können. Die Fülle und Unübersichtlichkeit des Materials kann jedoch die Auswahl geeigneter Videoclips für die Unterrichtsgestaltung erschweren (Brüning, 2018; Hogervorst, 2018, 2020).

Besuchen Zeitzeug:innen die Lernenden im Unterricht persönlich, stellt das für Schüler:innen eine besonders eindrückliche Form der Begegnung mit der Vergangenheit dar (Bertram, 2020; Henke-Bockschatz, 2007, 2014). Zeitzeug:innen besuchen die Schulen im Rahmen des Geschichtsunterrichts oder von Schulveranstaltungen, tragen ihre Erfahrungen vor und führen mit den Schüler:innen ein Gespräch, geben Antworten auf die persönlichen Fragen der Lernenden. Diese Form der Zeitzeug:innenbegegnung stellt für Schüler:innen und Lehrkräfte nicht nur eine Abwechslung zum üblichen Unterrichtsgeschehen dar, sondern ihr wohnt ein besonderes Motivationspotenzial inne (Bertram et al., 2017; Trofanenko, 2017). Zudem besitzt diese persönliche Begegnung „einen hohen Grad an Authentizität und Faszination“ (Bundesstiftung zur Aufarbeitung der SED-Diktatur, 2016). Trotz der Beliebtheit dieses Formats warnen Geschichtsdidaktiker:innen auch vor einem unreflektierten Einsatz der Berichte und fehlender Vor- und Nachbereitung (Bertram, 2017; Brauer, 2017; Henke-Bockschatz, 2007; von Plato, 2001): Schüler:innen fällt es häufig schwer, eine kritische Distanz zu wahren und die Berichte nicht als einzige Wahrheit zu empfinden. Die Begegnungen stellen somit hohe Anforderungen an das historische Lernen (Trofanenko, 2017).

In den vergangenen Jahren sind weitere Formate von Zeitzeug:innenberichten entstanden, die auf zwei Entwicklungslinien in der Gesellschaft zurückzuführen sind: (1) Mit dem zunehmenden zeitlichen Abstand zum Holocaust sterben die Überlebenden, die persönlich von dem Schrecken berichten können. Institutionen bemühen sich um eine Konservierung der Erfahrungsberichte. (2) Die Digitalisierung eröffnet neue Möglichkeiten in der Inszenierung und Konservierung der Zeitzeugnisse. Damit ergeben sich weitere Formate von Zeitzeug:innenberichten für das historische Lernen im Geschichtsunterricht: Beispielsweise werden derzeit vorwiegend in Museen oder in Gedenkstätten 3D-Zeugnisse von Holocaust Überlebenden entwickelt, die durch umfangreiche Interviews, entsprechende Programmierung und eine Spracherkennungssoftware lebensgroß in den Raum projiziert werden und mit Schüler:innen ins Gespräch kommen können (Ballis & Gloe, 2020). Weitere Bemühungen, persönliche Geschichten zu konservieren und für die jüngeren Generationen zugänglich zu machen, sind beispielsweise eine Augmented-Reality-App des Westdeutschen Rundfunks (WDR; 2021) und ein Instagram-Projekt des Südwestrundfunks (SWR; 2021). Mit der Augmented-Reality-App können Zeitzeug:innen als Hologramme in den Raum projiziert werden. Dabei sind die Holo-

gramme mit Musik und weiteren Spannungselementen (wie beispielsweise der Skyline des zerbombten London und Militärflugzeugen) inszeniert. Der SWR entwickelte das Instagram-Projekt @ichbinsophiescholl in Anlehnung an das israelische Instagram-Projekt @evastories (Henig & Ebbrecht-Hartmann, 2022). Beginnend ein Jahr vor ihrem Tod erzählt Sophie Scholl, von einer Schauspielerin gespielt, täglich in kurzen Videoclips von ihrem Alltag und antwortet auf Kommentare und Fragen der User:innen. Der SWR bietet für die Einbettung in den Unterricht eine ausgearbeitete Unterrichtsreihe mit Arbeitsblättern und Verlaufsplan an und verspricht: „Durch die intensive Beschäftigung mit der Biografie der jungen Frau erhalten die Schüler:innen facettenreiche Einblicke in Sophie Scholls Denken und Handeln“ (Zabori, 2021). Die neuen Projekte machen Lernenden Identifikations- bzw. Orientierungsangebote und definieren für sich einen Bildungsauftrag. Wenn auch öffentlich vielfach gelobt und teilweise ausgezeichnet, gab es auch Kritik an den neuen Formaten. Denn die Inszenierung setzt auf eine starke Emotionalisierung der Rezipient:innen und das Verhältnis von *wahren* Begebenheiten und Fiktion ist für die Rezipient:innen nur schwer zu erkennen (z. B. Dittmer, 2019; Meyer, 2021). Es wird deutlich, dass der Umfang an zugänglichen Erfahrungsberichten aus der Vergangenheit sowie die neuen Inszenierungen zusätzliche Herausforderungen an die pädagogische Arbeit mit den digitalen Zeitzeugnissen im Geschichtsunterricht stellen (Brüning, 2018; Hogervorst, 2020).

Wie genau der Einsatz von Zeitzeug:innenberichten auf die Schüler:innen wirkt, ob er emotionalisiert, Fehlvorstellungen von Geschichte vermittelt, motiviert oder Kompetenzen historischen Denkens fördert, sind wichtige empirische Fragen, die dabei helfen, ein gelungenes Lernarrangement mit Zeitzeug:innenberichten im schulischen Unterricht besser zu verstehen. Bisher gibt es allerdings nur wenige empirische Arbeiten, die untersuchen, wie das Lernen mit Zeitzeug:innenberichten gelingt und welche Auswirkungen das Lernarrangement mit Zeitzeug:innenberichten auf das Lernen haben. Im folgenden Kapitel stelle ich bisherige empirische Arbeiten zu dem Lernarrangement mit Zeitzeug:innenberichten vor. Dabei umfasst der Begriff des Lernarrangements verschiedene didaktische Ziele, Themen, Formate und Instruktionen.

### **1.2.3 Chancen und Risiken des Lernarrangements: Empirische Studien**

Während geschichtsdidaktische Handbücher und Handreichungen zum Einsatz von Zeitzeug:innenberichten im Geschichtsunterricht auf die Potenziale und Gefahren des Lernarrangements hinweisen (z. B. Ardemento & Kuszmar, 2013; Bundesstiftung zur Aufarbeitung der SED-Diktatur, 2016; Henke-Bockschatz, 2007), bleibt es eine empirische Frage, wie das Lernarrangement auf die Lernaktivität und das Lernergebnis wirkt. Bertram (2017) stellte in

ihrer Monografie zu Zeitzeug:innen im Geschichtsunterricht die bis dato durchgeführten empirischen (und nach Bertram belastbaren) Studien zur Wirksamkeit von Zeitzeug:innenarbeit in der Schule zusammen. Sie folgert aus den vier empirischen Studien (Galda, 2013; Lanman, 1987; Obens & Geißler-Jagodzinski, 2010; Rodenhäuser, 2012), dass es Lernenden aufgrund ihrer emotionalen Ergriffenheit häufig schwerfällt, die Berichte kritisch zu hinterfragen, dass aber, so zeige die Studie von Rodenhäuser, der kritische Umgang mit entsprechender Unterstützung gelingen kann. Mit ihrer randomisierten Interventionsstudie begegneten Bertram et al. (2017) der schwachen empirischen Forschungslage zur Wirksamkeit von Zeitzeug:innenberichten im Geschichtsunterricht. Die Ergebnisse der Studie bestätigten die bisherigen Forschungsergebnisse mit einer großen Stichprobe ( $N = 962$ ) und mit einem randomisiert-kontrollierten Studiendesign: Schüler:innen, die mit Zeitzeug:innen live, mit videografierten oder transkribierten Interviewsequenzen arbeiteten, lernten mit Blick auf Einsichten in zentrale epistemologische Prinzipien von Geschichte mehr als eine Kontrollgruppe, die nicht mit Zeitzeug:innenberichten arbeitete, sondern dem regulären Geschichtsunterricht mit ihrer Geschichtslehrkraft folgte. Eine differenziertere Betrachtung der einzelnen Interventionsbedingungen zeigte jedoch, dass Schüler:innen in der Live-Bedingung im Vergleich zu Mitschüler:innen in den anderen Studienbedingungen zwar mehr Freude an der Unterrichtseinheit hatten und überzeugter davon waren, etwas aus der Unterrichtsreihe gelernt zu haben, allerdings weniger Einsicht in den Konstruktcharakter von Geschichte zeigten. Dies bedeutet, dass der Unterricht mit Live-Zeitzeug:innen die Lernenden motiviert und ihnen das Gefühl vermittelt, die Vergangenheit durch die Berichte besser verstehen zu können, obwohl der Unterricht mit Live-Zeitzeug:innen Kompetenzen historischen Denkens weniger erfolgreich förderte als der Unterricht mit konservierten Zeitzeug:innenberichten.

In den vergangenen Jahren gab es weitere empirische Bemühungen, das Lernen mit Zeitzeug:innen im Geschichtsunterricht zu verstehen.<sup>8</sup> Brüning (2018) stellte in ihrer Arbeit die Frage, was Schüler:innen in heterogenen Lerngruppen in Projekttagen mit videografierten Berichten Holocaust-Überlebender aus dem *Visual History Archive* lernen. Das Erkenntnisinteresse ihrer Forschungsarbeit war primär deskriptiver Natur. Brüning setzte einen standardisierten Fragebogen ein, führte Einzel- und Gruppeninterviews sowie eine abschließende Gruppendiskussion mit 122 Schüler:innen aus fünf Lerngruppen. Wie in bisherigen Forschungsarbeiten zeigten sich auch die Schüler:innen in Brünings Studie beeindruckt von den authentischen Geschichten: „Es war als würde man Geschichte lebendig erleben. Ich würde

---

<sup>8</sup> In der Zusammenstellung der Studien, die hier aufgeführt werden, fanden einzelne Fallstudien (z. B. Dutt-Doner, Allen & Campanaro, 2016) keine Berücksichtigung.

sagen ich habe Geschichte live erlebt“ (Schülerzitat aus Brüning, 2018, S. 320). Die Autorin zog das Fazit, dass das Lernen mit den authentischen Geschichten Erfahrungen wie Immersion und Empathie initiiert, eine kritische Analyse der Berichte diesen Aspekten allerdings zuwiderläuft. Brüning forderte die Entwicklung einer Art der Quellenkritik, die dem Medium (digitalisierte Zeitzeugnisse) gerecht wird. Auch Hogervorst (2020) interessierte sich für videografierte Zeugnisse: Sie befragte 167 Nutzer:innen der niederländischen Plattform [www.getuigenverhalen.nl](http://www.getuigenverhalen.nl), die etwa 500 Interviews mit Überlebenden des Holocaust und des Zweiten Weltkriegs aufbereitet hat. Die Nutzungsstatistiken zeigten, dass die Website durchschnittlich 2 100-mal im Monat aufgerufen wurde, wobei allerdings niemand ein Interview – die Interviews dauern etwa 90 Minuten – bis zum Ende anschaute. Die längste Verweildauer auf der Seite betrug 29 Minuten. Hogervorst schlussfolgerte mit Blick auf die Umfrage der Nutzer:innen, dass trotz der Distanz eine Art emotionales Involvement entsteht. Auch in diesem Setting zeigte sich, dass die Nutzer:innen den Berichten aufgrund der persönlichen Erfahrungen in der Vergangenheit eine besondere Bedeutung in der Vermittlung von Geschichte zuschrieben.

Jüngst veröffentlichte Richardson (2021) Untersuchungen zu dem britischen Bildungsprogramm *Lessons from Auschwitz*, in dem Schüler:innen eine:n Holocaustüberlebende:n treffen, 60 Minuten ihren:seinen Erzählungen zuhören und dann 30 Minuten lang Fragen stellen. Richardson befragte 58 Teilnehmende in einer Online-Umfrage und strukturierten Interviews, welche Gedanken die Teilnehmenden zu der Begegnung mit den Zeitzeug:innen hatten. Aus den Antworten gingen drei Merkmale der Begegnungen hervor: *concrete*, *connecting*, *current*. *Concrete* meint die Besonderheit, dass die Begegnung aktiver war als herkömmliche Arbeit im Geschichtsunterricht. *Connecting* umfasst die Verbindung zum eigenen Leben durch die Erzählung von alltäglichen Routinen oder Details zur Familie. Die Eigenschaft *current* bezieht sich auf das Erkennen der Aktualität der Erzählung und der persönlichen Verantwortung der Lernenden in ihrem persönlichen Handeln. Richardson warf zudem die Frage auf, ob es wünschenswert sei, dass Schüler:innen die Erfahrungen der Holocaustüberlebenden verstehen, und was sie damit meinten, wenn sie von „verstehen“ sprachen: Können sie die Individualität eines jeden Zeugnisses verstehen? Richardson folgerte: „We would perhaps do our students a better service if we concentrated our efforts more on helping them see the testimony as particular, but unexceptional; an individual’s story, but not the *only* story – and certainly not the *whole* story“ (S. 17).

Ballis und Schwendemann (2021) untersuchten aus sprachwissenschaftlicher Perspektive mit einem Mixed-Methods-Ansatz die kommunikative Begegnung von Schüler:innen mit einem Holocaust-Überlebenden. 575 Schüler:innen nahmen an der quantitativen Befragung teil,

97 zusätzlich in qualitativen Interviews. Die Schüler:innen waren überzeugt, dass sie durch die Begegnung mit der *first-hand source* etwas über die Vergangenheit lernen konnten. Die Autorinnen stellten heraus, dass das Vertrauen der Schüler:innen gegenüber den Zeitzeug:innenberichten das Fundament der Begegnungen ist.

Trotz der Unterschiedlichkeit der verschiedenen empirischen Studien mit Blick auf das Format der Zeitzeug:innenberichte, des methodischen sowie fachwissenschaftlichen Zugangs und des konkreten Fokus der Fragestellungen bestätigen die empirischen Studien zur Arbeit mit Zeitzeug:innenberichten im formalen Bildungskontext in der Gesamtschau die Chancen und Risiken des Lernarrangements: Schüler:innen haben Interesse an der Art und Weise, wie Geschichte erzählt wird, schenken den Berichten allerdings ein sehr starkes Vertrauen, wahre Auskünfte über die Vergangenheit zu geben: „[...] in any case, you believe him one hundred percent, everything he says“ (Schülerin in Ballis & Schwendemann, 2021, S. 12). Entgegen der Studie von Bertram et al. (2017) geben die Studien von Brüning (2018) und Hogervorst (2020) Hinweise darauf, dass auch videografierte Zeitzeugnisse einen Überwältigungscharakter haben können, und die Studie von Bertram et al. (2017) zeigt wiederum, dass der Überwältigungscharakter nicht auf die Begegnung mit Holocaust-Überlebenden beschränkt ist. Allerdings ist zu berücksichtigen, dass die empirischen Studien bis auf die Studie von Bertram et al. (2017) keine kausalen Schlüsse (im Sinne von: Begegnung führt zu Überwältigung) zulassen. Das starke Vertrauen, das Lernende den Berichten entgegenbringen, *die Wahrheit* über die Vergangenheit zu erfahren, steht dem Ziel des Geschichtsunterrichts, Kompetenzen historischen Denkens zu fördern und ein reflexives Geschichtsbewusstsein auszubilden, entgegen.

Es stellt sich die Frage, wie das Lernarrangement gestaltet sein muss, damit Schüler:innen das Lernziel erreichen. Die pädagogisch-psychologische Unterrichtsforschung gibt dazu Hinweise, weshalb im folgenden Kapitel das Lernarrangement durch die Unterscheidung zwischen Oberflächen- und Tiefenmerkmalen betrachtet wird. Zur Verknüpfung der generischen Tiefenmerkmale mit fachspezifischen Merkmalen des Geschichtsunterrichts gibt es zum jetzigen Zeitpunkt lediglich theoretische und konzeptuelle Überlegungen (Trautwein et al., 2021; Zülsdorf-Kersting, 2020), obwohl ein solches Zusammenbringen von allgemeiner Unterrichtsforschung und Geschichtsdidaktik für beide Seiten einen „hohen Erkenntnisgewinn“ (Trautwein et al., 2021, S. 271) verspricht.

### **1.2.4 Oberflächen- und Tiefenmerkmale des Lernarrangements mit Zeitzeug:innenberichten**

„Doch braucht das Lernen über Geschichte wirklich Zeitzeugen? Die Antwort ist ein klares ‚Ja!, aber‘. Denn es gilt, einiges dabei zu bedenken“ (Brauer, 2017). Da die Unterrichtsforschung sich unter anderem mit der Frage beschäftigt, was qualitativollen Unterricht ausmacht (siehe Kapitel 1.1.2), kann sie Hinweise darauf liefern, was es beim Lernarrangement mit Zeitzeug:innen zu bedenken gibt, damit es das historische Lernen fördern kann. Eine Verknüpfung der Oberflächen- und Tiefenmerkmale mit den empirischen Studien zum Lernen mit Zeitzeug:innenberichten (Kapitel 1.2.3) kann in Anlehnung an Trautwein et al. (2021) Hinweise für Gelingensfaktoren für das Lernarrangement mit Zeitzeug:innenberichten aufzeigen.

Entsprechend den empirischen Ergebnissen der Unterrichtsforschung zu den Oberflächen- und Tiefenstrukturen geben die empirischen Studien zu Zeitzeug:innenberichten im Unterricht Hinweise darauf, dass der methodische Zugang zu den Berichten keine Bedeutung für das Lernen hat. Ob Projekttag (Brüning, 2018), Interviews (Bertram et al., 2017) oder Vorträge (Ballis & Schwendemann, 2021; Richardson, 2021), die Studien zogen alle das Fazit, dass die Lernenden den Berichten ein besonderes Vertrauen entgegenbringen, wahre Einblicke in die Vergangenheit zu geben. Entgegen der Annahme und Befunde, die Oberflächenmerkmale spielten eine weniger bedeutsame Rolle im Unterrichtsgeschehen, weist die Studie von Bertram et al. (2017) allerdings darauf hin, dass das Format des Berichts – Live, Video oder Text – eine Bedeutung für das Lernen hat. Dieser Befund liefert erste Hinweise, dass im Lernarrangement mit Zeitzeug:innenberichten auch die Oberflächenmerkmale relevant sein könnten.

Darüber hinaus können die Tiefenmerkmale Anknüpfungspunkte dafür geben, wie das Lernen mit Zeitzeug:innenberichten zu einem gelungenen Lernarrangement gestaltet werden kann. Das Tiefenmerkmal der konstruktiven Unterstützung umfasst methodisch-didaktische sowie sozio-emotionale Anteile von Unterstützung (Trautwein et al., 2021) und beschreibt den wertschätzenden Umgang zwischen Lernenden und Lehrkraft. Der aktuelle Forschungsstand zum Lernen mit Zeitzeug:innenberichten zeichnet ein konsistentes Bild: Die Berichte berühren viele Schüler:innen emotional, insbesondere dann, wenn die Zeitzeug:innen Gewalt und Unrecht erfahren haben (Ballis & Schwendemann, 2021; Brüning, 2018; Obens & Geißler-Jagodzinski, 2010). Die emotionale Ergriffenheit der Schüler:innen erfordert die sozio-emotionale Unterstützung durch die Lehrkraft, die im Sinne einer „Kultivierung der Affekte“ (Mitscherlich, 1983, S. 40, zitiert nach von Borries, 1994) die Gefühle der Schüler:innen anerkennt und diesen Raum gibt, sie zu thematisieren. Ein fehlendes Thematisieren der Gefühle der Schüler:innen kann dazu führen, dass diese die Gefühle der Zeitzeug:innen nachempfinden

(Richardson, 2021), was dem Beutelsbacher Konsens (siehe Kapitel 1.1.3, Wehling, 1977) widerspricht.

Der kognitive Lernprozess wird vom Tiefenmerkmal Potenzial zur kognitiven Aktivierung im Modell der drei Basisdimensionen abgebildet. Das Merkmal beschreibt die Tiefe der Lernprozesse durch anregende Aufgaben und Unterrichtssituationen wie beispielsweise kognitive Konflikte. Kognitive Aktivierung sei „unabdingbar, wenn Orientierungskompetenz jenseits der kritiklosen Übernahme von Orientierungsangeboten angestrebt“ (Trautwein et al., 2021, S. 267) würde. Die empirischen Studien liefern Hinweise darauf, dass es Schüler:innen schwerfällt, einen Zeitzeug:innenbericht zu dekonstruieren und mit dem Konstruktcharakter von Geschichte in Verbindung zu bringen (z. B. Bertram et al., 2017; Brüning, 2018). Rodenhäuser (2012) fand in ihrer Studie allerdings erste Hinweise, dass mit einer strukturierten Anleitung eine kritische Reflexion der Begegnung möglich sei. Deshalb erscheint die methodisch-didaktische Einbettung der Begegnungen als besonders bedeutsam. Entsprechend zog Brüning (2019) in ihrer Studie das Fazit, dass es für den reflektierten Umgang mit Zeitzeug:innenberichten eine besondere Art der Quellenkritik benötige, um kognitive Prozesse im Sinne des historischen Denkprozesses anzustoßen und sowohl eine sachbezogene Ebene sowie eine metakognitiv-epistemologische Ebene zu bedienen (Zülsdorf-Kersting, 2020).

Das Tiefenmerkmal Klassenführung beschreibt, wie gut der Unterricht von der Lehrkraft gesteuert ist und Ziele klar formuliert werden, damit alle Schüler:innen möglichst viel Lernzeit effektiv nutzen können. Das Merkmal ist „relevant bei der Realisierung offener, kommunikativer, nicht-linearer Unterrichtssituationen“ (Trautwein et al., 2021, S. 268), wie es bei der – vor allem persönlichen – Begegnung mit Zeitzeug:innenberichten der Fall ist. Die Begegnung mit Live-Zeitzeug:innen stellt eine kommunikative Situation dar, die eine klare Moderation erfordert (Ballis & Schwendemann, 2021).

Guter Geschichtsunterricht, der zu Orientierung führt, könne nur „bei einer gelungenen Synthese von kognitiver Aktivierung und sozio-emotionaler-konstruktiver Unterstützung“ (S. 269) gelingen, so Trautwein et al. (2021). Die Beschreibung des Lernarrangements anhand der Tiefenmerkmale verdeutlicht, dass hier die Basisdimensionen sehr eng miteinander verbunden sind, und gibt Anknüpfungspunkte für die Gelingensfaktoren des Lernarrangements mit Zeitzeug:innenberichten, wengleich die aufgeführten Studien zu Zeitzeug:innenberichten nicht explizit die Tiefenmerkmale adressieren. Der historische Denkprozess und ein reflexives Geschichtsbewusstsein können mit Zeitzeug:innenberichten nur dann initiiert und gefördert werden, wenn die emotionale Betroffenheit der Schüler:innen durch die konstruktive



---

Unterstützung der Lehrkraft anerkannt und thematisiert wird und in eine kognitiv-aktivierende Vor- und Nachbereitung eingebettet ist. Das kann nur gelingen, wenn trotz kommunikativer und nicht-linearer Unterrichtssituation das Unterrichtsgeschehen klar mit Blick auf das didaktische Ziel geführt wird. In den Ausführungen wird deutlich, dass in den Tiefenmerkmalen des Unterrichts die Lehr-Lern- und Interaktionsprozesse abgebildet sind, die zwar häufig auf Angebotsseite des Angebots-Nutzungs-Modells verortet wurden, aber vielmehr eine Verknüpfung von Lehr- und Lernprozessen darstellen, d. h. Angebot und Nutzung miteinander verbinden (Seidel, 2014). Das Angebot-Nutzungs-Modell nimmt an, dass das unterrichtliche Angebot nicht unmittelbar, sondern über die Lernaktivität der Schüler:innen auf den Kompetenzerwerb wirkt. Das bedeutet, dass eine Berücksichtigung der Faktoren aufseiten der Lernenden zentral erscheint, um ein umfangreiches Verständnis für das Lernarrangement mit Zeitzeug:innenberichten zu erhalten. Diese standen allerdings nur selten im Fokus bisheriger Forschungsarbeiten. Deshalb thematisiere ich im folgenden Kapitel die Faktoren aufseiten der Schüler:innen, ihr Lernpotenzial, ihre Wahrnehmung und Interpretation und schließlich die Nutzung des Unterrichtsangebots mit Zeitzeug:innenberichten: die Lernaktivität.

### **1.3 Lernpotenzial, Wahrnehmung und Lernaktivität: Lernende begegnen Zeitzeug:innenberichten**

Die bisherigen Ausführungen haben gezeigt, dass theoretische und empirische Studien den Zeitzeug:innenberichten eine besondere Wirkmächtigkeit attestieren, die sich in einem starken Vertrauen der Lernenden äußert, dass diese Berichte *wahre* Einblicke in die Vergangenheit geben würden. Diese Fehlvorstellungen von Geschichte stehen in Konflikt mit den Zielen des Geschichtsunterrichts. Dem Angebots-Nutzungs-Modell für die Wirkungsweisen von Unterricht liegt die Annahme zugrunde, dass das unterrichtliche Angebot nicht unmittelbar auf das Lernergebnis wirkt, sondern auch von Faktoren aufseiten der Schüler:innen abhängt, von ihrem Lernpotenzial, ihrer Nutzung des Unterrichtsangebots – der Lernaktivität – und der Wahrnehmung des Unterrichts. Zudem berücksichtigt das Modell den familiären Hintergrund als Einflussfaktor auf die Lernprozesse. Um das Lernarrangement mit Zeitzeug:innenberichten und seine Wirkung besser zu verstehen, werde ich im folgenden Kapitel zunächst auf die Bedeutung der Lernvoraussetzungen von Schüler:innen im Geschichtsunterricht mit Zeitzeug:innenberichten eingehen, ehe ich die Interpretation und Wahrnehmung des Lernmaterials in Form der (wahrgenommenen) Authentizität thematisiere und schließlich ein Modell mit vier Verarbeitungsmodi entwickle, das die Verarbeitungsprozesse im Rahmen der Lernaktivität mit Zeitzeug:innenberichten beschreibt.

#### **1.3.1 Individuelle Lernvoraussetzungen und ihr Zusammenspiel mit der Unterrichtsgestaltung**

Lernende bringen auf verschiedenen Ebenen unterschiedliche Lernvoraussetzungen mit. Diese Heterogenität betrifft beispielsweise kognitive, affektive und motivationale Lernvoraussetzungen (Helmke, 2017; Vieluf et al., 2020). Lernen im Sinne der Wissens- und Kompetenzzaneignung ist immer eine Verknüpfung mit bereits bestehendem Wissen (Renkl, 2015). Dieses Vorwissen stellt eine kognitive Lernvoraussetzung dar: Neue Informationen werden dann im Arbeitsgedächtnis verarbeitet, wenn sie als relevant interpretiert werden. Das hängt maßgeblich vom Vorwissen ab. Der Einfluss des Vorwissens auf die Leistung von Schüler:innen wurde vielfach erforscht und gilt mittlerweile als *Common Sense* in der Pädagogischen Psychologie (z. B. Fyfe & Rittle-Johnson, 2016; Hampp & Schwan, 2017; Renkl, 2015; Simonsmeier, Flaig, Deiglmayr, Schalk & Schneider, 2022). Auch im Kontext historischen Lernens weisen die Modelle historischen Denkens und verwandte Konzepte über die Operation der Kontextualisierung (z. B. Endacott & Brooks, 2013; Van Drie & Van Boxtel, 2008; Wineburg, 1991) auf die Bedeutung des Vorwissens hin: Ohne Wissen über die Zeit, von der

die Zeitzeug:innen berichten, können die Berichte weder angemessen verarbeitet werden noch ist es möglich, historische Empathie zu empfinden (z. B. Endacott & Brooks, 2013).

Auch kognitive und meta-kognitive Lernstrategien und Überzeugungen zählen zum kognitiven Lernpotenzial. Hierunter lassen sich mit Blick auf das historische Lernen mit Zeitzeug:innenberichten epistemologische Überzeugungen von Geschichte fassen (Bertram, Wagner & Trautwein, 2014; Maggioni, Alexander & VanSledright, 2004; Maggioni, VanSledright & Alexander, 2009). Wissen Lernende, dass es einen Unterschied zwischen Vergangenheit und Geschichte gibt, dass jede:r Zeitzeug:in eine individuelle Sicht auf die Vergangenheit hat und folglich niemals eine Perspektive als *die Wahrheit* bezeichnet werden kann und darf? Diese Einsichten sind notwendig, um die Berichte angemessen zu dekonstruieren und Inkonsistenzen zwischen Perspektiven zu erkennen. Brüning (2018) stellte fest, dass den Schüler:innen das nötige Vorwissen und die Konzepte im Umgang mit der besonderen Quellengattung fehlten.

Die Bedeutung von Motivation und Interesse für den Lernerfolg ist unumstritten (z. B. Dweck, 1986; Harackiewicz, Smith & Priniski, 2016; Krapp, 2002; Schiefele, 1991) und mittlerweile vielfach empirisch belegt (z. B. Gaspard et al., 2015; Hampf & Schwan, 2017; Hulleman, Barron, Kosovich & Lazowski, 2016; Hulleman & Harackiewicz, 2009; Schiefele, Krapp & Winteler, 1992). Nach einer Umfrage der Körber-Stiftung (2017) zeigten 56 % der befragten Jugendlichen ein sehr großes bzw. großes Interesse an Geschichte. Wenn Schüler:innen allerdings über die Nützlichkeit des Lernens über Geschichte befragt wurden, zeigte sich, dass nur wenige Schüler:innen die Verbindung zur Gegenwart herstellen können, folglich die Relevanz des historischen Lernens nicht erkennen (Haydn & Harris, 2010; Lee, 2004; Van Straaten et al., 2016).

Durch den individuellen Zugang verfügen Zeitzeug:innenberichte über das Potenzial, das Interesse der Lernenden zu wecken (Bertram et al., 2017). Denn die Zeitzeug:innenberichte füllen die abstrakten historischen Ereignisse mit lebendigen Erfahrungen und verknüpfen die Ereignisse häufig mit einer Botschaft für die Gegenwart. Schließlich unterscheiden sich Lernende auch mit Blick auf affektive Komponenten des Lernpotenzials. So können beispielsweise persönliche Erfahrungen von Ausgrenzung oder Flucht mit den Berichten der Zeitzeug:innen verglichen werden. In diesen Fällen erscheint es offensichtlich, dass Lernende mit vergleichbaren Erfahrungen anders angesprochen werden als Lernende, die keine Parallelen zu ihrem Leben sehen. Außerdem ist auch die Neigung zu bestimmten Gefühlen wie Ängstlichkeit und Freude von Bedeutung für die Lernaktivität. Bilandzic, Schnell, Sukalla, Hastall and Busselle

(2019) entwickelten das Konzept der *Narrative Engageability*, einer stabilen Persönlichkeitseigenschaft, die die Tendenz zum Eintauchen in eine Geschichte beschreibt. Mit Blick auf das Lernen mit Zeitzeug:innenberichten ist folglich anzunehmen, dass derartige Eigenschaften bei den Lernenden unterschiedlich ausgeprägt sind und die Verarbeitung des Lernmaterials entsprechend unterschiedlich beeinflussen.

Zahlreiche Studien können mittlerweile die Bedeutung der Familie für die Lernprozesse und Bildungsabschlüsse bestätigen: Diese Bedeutung reicht von strukturellen Merkmalen der Familie wie dem sozioökonomischen Status (z. B. OECD, 2005; White, 1982) oder dem Migrationshintergrund (z. B. Fuligni, 1997; Potter & Morris, 2017; Roscigno & Ainsworth-Darnell, 1999) bis zu Prozessmerkmalen wie der Motivation der Eltern für den Lernerfolg ihrer Kinder (z. B. Higgins & Katsipataki, 2015) und ihre Involviertheit bei schulischen Aufgaben (z. B. Häfner et al., 2018). Im Geschichtsunterricht zur Zeitgeschichte und insbesondere auch im Lernarrangement mit Zeitzeug:innenberichten kommt dem familiären Bezug zur Thematik eine besondere Rolle zu, denn er kann einen Einfluss darauf haben, wie involviert die Lernenden sind (Green, 2004; Hofer, 2016). Dieser familiäre Bezug äußert sich auf verschiedene Weise. Zwei Möglichkeiten werden hier beispielhaft aufgeführt: (1) Jede Familie hat eine Vergangenheit und insbesondere zur Zeitgeschichte persönliche Bezüge. Familien entwickeln Narrative über die Familienbiografie und geben diese an die nachfolgenden Generationen weiter (Rosenzweig & Thelen, 1998). Lernende können so in ihrer Familie selbst erfragen, wie die Eltern oder Großeltern etwa den Fall der Mauer erlebt haben und Informationen der Zeitzeug:innenberichte mit Familiennarrativen vergleichen: War der Fall der Mauer wirklich ein so einschneidendes Erlebnis? Familiennarrative sind allerdings häufig „aufgehübscht“ und beinhalten eher Ereignisse, die einem positiven Selbstbild der Familie förderlich sind (Erl, 2011; Gudjons, Wagener-Gudjons & Pieper, 2008). Sie tragen folglich nicht unbedingt zu einer reflektierten Einordnung der Perspektiven bei. (2) Lernende mit unterschiedlicher ethnischer und religiöser Zugehörigkeit unterscheiden sich in ihrem Bezug zu den historischen Themen. Wie reagiert beispielsweise ein palästinensisches Kind, dessen Großeltern durch die Siedlungspolitik Israels vertrieben wurde, auf eine:n Zeitzeug:in, der:die im Zweiten Weltkrieg nach Israel floh und dort nun seit Jahren lebt? (Goldberg, 2014; Heafner & Fitchett, 2015; Trautwein et al., 2021). Auch bei anderen Konflikten in der Weltgeschichte sind direkte familiäre Bezüge der Lernenden denkbar, die sie in ihren Lernprozessen beeinflussen (z. B. Konflikt zwischen Kurden und Türken; Völkermord an Armeniern; ganz aktuell: Russlands Angriffskrieg gegen die Ukraine). Brüning (2019) konnte in ihrer Untersuchung zum Einsatz videografiert Zeitzeugnisse in heterogenen Klassen (siehe Kapitel 1.2.3) im Umgang mit den Berichten keine

Unterschiede zwischen Schüler:innen deutscher und nicht deutscher Herkunft feststellen. Knapp die Hälfte der Stichprobe von fünf Schulklassen gab im Fragebogen an, einen Migrationshintergrund zu haben. Brüning untersuchte allerdings nicht, ob die Schüler:innen sich in ihrem familiären Bezug zum Zweiten Weltkrieg unterschieden.

Die Heterogenität von Lerngruppen zeigt sich folglich auf vielen verschiedenen Ebenen. Sie hat nicht nur direkten Einfluss auf die Lernprozesse der Schüler:innen, sondern ist auch entscheidend dafür, wie ein spezifisches Lernarrangements wirkt. Welche Wirkung das Unterrichtsangebot auf Lernende mit unterschiedlichen Lernvoraussetzungen hat, ist Forschungsgegenstand verschiedener Forschungslinien. Studien zu adaptivem Unterricht untersuchen beispielsweise, wie Lehrkräfte auf die individuellen Unterschiede ihrer Schüler:innen reagieren (Paas, van Merriënboer & van Gog, 2012; Parsons et al., 2017) und wie adaptives Feedback in Online-Lernumgebungen das Lernen der Schüler:innen fördern kann (Ma, Adesope, Nesbit & Liu, 2014). Forschungsarbeiten im Rahmen des Aptitude-Treatment-Interaction-Paradigmas (ATI; Cronbach, 1957) beschäftigen sich mit der Frage, unter welchen Bedingungen der Unterricht für Schüler:innengruppen mit bestimmten Voraussetzungen förderlich ist. Empirische Studien fanden beispielsweise Hinweise darauf, dass die Gestaltung des Lernmaterials einen Einfluss auf das Lernen in Abhängigkeit der Arbeitsspeicherkapazität (z. B. Fuchs et al., 2014; Lehmann, Goussios & Seufert, 2016), der Präferenzen der Lernenden (Lehmann & Seufert, 2020) oder des Autonomieerlebens (Flunger, Mayer & Umbach, 2019) hat. Diese Forschungsarbeiten weisen mit Blick auf das Lernarrangement mit Zeitzeug:innenberichten darauf hin, dass die Gestaltung des Lernarrangements (z. B. durch Format, Inhalt, Zielsetzungen) in Abhängigkeit der familiären Hintergründe sowie der kognitiven, motivationalen und affektiven Voraussetzungen der Lernenden auf die Lernaktivität und das Lernergebnis der Schüler:innen wirkt. Für die Wirkung des Unterrichts sind allerdings nicht nur der familiäre Hintergrund und das individuelle Lernpotenzial von Bedeutung, sondern auch, wie das Unterrichtsangebot durch die Lernenden wahrgenommen und interpretiert wird. Insbesondere beim Lernarrangement mit Zeitzeug:innenberichten spielt die Zuschreibung von Authentizität eine wichtige Rolle für die Lernaktivität und soll im folgenden Abschnitt thematisiert werden.

### **1.3.2 Wahrnehmung und Interpretation: Authentizität von Zeitzeug:innenberichten**

„Anschaulich und authentisch“, „living history“, „aus erster Hand“ – so betitelten Lehrkräfte und ihre Schüler:innen die Erfahrung mit dem Besuch eines:r Zeitzeug:in aus der ehe-

maligen DDR (<https://www.ddd-zeitzeuge.de/ddd-zeitzeugen-gespraech.html>). Die Zeitzeug:innenberichte konfrontieren Schüler:innen mit „authentischer Geschichte“ (Sauer, 2006, S. 238) und ihre Wirkmächtigkeit wird häufig durch eine „Aura der Authentizität“ (Ballis & Schwendemann, 2021; Bertram, 2017; Brüning, 2018; Sabrow, 2012) erklärt. Der Terminus „Authentizität“ ist auf die griechischen Worte *authentēs* (Urheber:in, Täter:in, Mörder:in, Ausführer:in) und *authentikos* (zuverlässig, echt original, richtig; Ueding, 2012) zurückzuführen. Das Adjektiv „authentisch“ wurde in der mittelalterlichen Auseinandersetzung mit der Bibel als synonym für „wahr“ verwendet (Ueding, 2012). Authentizität ist Gegenstand verschiedener Forschungsdisziplinen wie der Anthropologie (z. B. Handler, 1986; Handler & Saxton, 1988), der Philosophie (z. B. Guignon, 2008), der Museumsforschung (z. B. Frazier, Gelman, Wilson & Hood, 2009; Schwan & Dutz, 2020), der Rhetorik (z. B. Ostermann, 2002; Ueding, 2012), der Pädagogischen Psychologie (z. B. Fougat, Misfeldt & Shaffer, 2019; Shaffer & Resnick, 1999) oder den Geschichtswissenschaften (z. B. Saupe, 2017). Die diversen Zugänge sind Grund dafür, dass es keine allgemeingültige Definition des Begriffs gibt (z. B. Krämer, 2012; Saupe, 2017). Häufig wird das Konzept über verschiedene Dimensionen beschrieben, die teilweise über Disziplinen hinweg vergleichbar sind. In Anlehnung an die Rhetorik wird auch in der Geschichtswissenschaft eine Unterscheidung zwischen persönlicher und materieller Authentizität vorgenommen (z. B. Krämer, 2012; Ueding, 2012). Während persönliche Authentizität eine Konsistenz zwischen Innenleben der Person und der Erzählung umfasst, folglich die Glaubwürdigkeit einer Person bezeichnet, bezieht sich die materielle Authentizität auf die Urheberschaft und Echtheit von Objekten, wie beispielsweise Dokumenten. Authentizität kann sich also auf Subjekte und Objekte beziehen. In der Rhetorik umfasst die persönliche Authentizität nicht nur die Konsistenz zwischen Innenleben und Erzählung, sondern auch, wie logisch und nachvollziehbar eine Erzählung an sich ist (Ueding, 2012). Diese Form der Authentizität spielt auch in der Medienpsychologie eine Rolle und wird dort als narrative Authentizität bezeichnet (Busselle & Bilandzic, 2008; Hofer, 2016; Weber, 2021). Die Besonderheit von Zeitzeug:innenberichten liegt unter anderem darin, dass sie beide Dimensionen vereinen: Für die Authentizität des Subjekts ist es von Bedeutung, wie ehrlich und glaubwürdig die Zeitzeug:innen sich den Lernenden öffnen und ob die Lernenden der Geschichte folgen können: Spiegelt die Erzählung das Innenleben der Person wider? Ist die Erzählung in sich schlüssig? Gleichzeitig stellen Zeitzeug:innenberichte als „gegenständliches Relikt“ (Krämer, 2012, S. 22) eines vergangenen Ereignisses eine Quelle der Vergangenheit dar; sie waren Zeug:innen einer bestimmten historischen Zeit oder eines Ereignisses. Diese beiden Dimensionen beziehen sich

auf den empirischen Authentizitätsbegriff (quellenkritisch gesicherte Urheberschaft) und den normativen Authentizitätsbegriff im Sinne des Sich-selbst-treu-Seins (Ueding, 2012).

Der interpretative Authentizitätsbegriff stellt eine weitere Form des Verständnisses von Authentizität dar, nach der die Authentizität das Resultat eines Zuschreibungsprozesses ist, folglich nicht nur von den Zeitzeug:innen selbst, sondern insbesondere von den Lernenden (Ueding, 2012) und von einer Interaktion zwischen Lernenden und Zeitzeug:innenbericht (Kirchberg & Tröndle, 2012; Schwan & Dutz, 2020) ausgeht. Dieses Verständnis von Authentizität ist beim Lernen mit Zeitzeug:innenberichten zentral und findet sich im Angebots-Nutzungs-Modell unter dem Aspekt „Wahrnehmung und Interpretation“ des Unterrichtsangebots wieder (Helmke 2017). Auch im Rahmen dieses interpretativen Authentizitätsbegriffs lassen sich verschiedene Dimensionen finden, die die Authentizität eines Objekts, einer Person oder Erzählung konstituieren: Hampp und Schwan fanden in ihren Umfragen von Museumsbesucher:innen beispielsweise heraus, dass Teilnehmende unterschiedliche Begründungen lieferten, weshalb sie ein Objekt als authentisch wahrnahmen. Aus diesen Begründungen arbeiteten die Autor:innen die folgenden Dimensionen heraus: Geschichte, Charisma, Seltenheit, Prestige und Funktionalität/Vollständigkeit (Hampp & Schwan, 2014a, 2014b, 2017; Schwan & Dutz, 2020). Auch Hofer (2016) entwickelte in Anlehnung an verschiedene medienpsychologische Studien eine Klassifizierung wahrgenommener Authentizität von Medieninhalten mit den folgenden Dimensionen: Faktualität, Plausibilität, Typizität, narrative Konsistenz und wahrgenommene Nützlichkeit.

In der Wahrnehmung von Zeitzeug:innenberichten als „authentisch“ zeigt sich das große Vertrauen, das Lernende dieser Darstellungsform entgegenbringen (z. B. Angvik & von Borries, 1997; Krämer, 2012; Rosenzweig & Thelen, 1998). Sie trauen Zeitzeug:innenberichten sogar mehr als Lehrkräften oder Schulbuchttexten zu, etwas über die Vergangenheit lehren zu können (Angvik & von Borries, 1997). Dieser Vertrauensvorsprung weist darauf hin, dass mitunter Fehlvorstellungen von Geschichte vorliegen. Denn Zeitzeug:innen können zwar ein historisches Ereignis bezeugen (im Sinne der empirischen Authentizität) und nach bestem Gewissen ehrlich über ihre Gefühle berichten (im Sinne der normativen Authentizität), dennoch verändert sich die Erinnerung der Zeitzeug:innen mit der Zeit: Zum einen wird bestimmten Details eine höhere Bedeutung beigemessen als anderen, zum anderen nehmen sie in der Erinnerung und Erzählung selbst eine Deutung der historischen Ereignisse vor, die auch falsch sein könnte (Krämer, 2012). Die „Aura der Authentizität“, so Heyl (1995), kann eine Wahrnehmungsblockade schaffen. Nehmen Lernende die Berichte als authentisch wahr und sind sie in die Geschichten involviert, fehlt es möglicherweise an kognitiven Kapazitäten, den Bericht auf

einer abstrakten Ebene im Sinne des historischen Denkprozesses zu verstehen und die Authentizität zu hinterfragen: „People are credulous creatures who find it very easy to believe and very difficult to doubt“ (Gilbert, 1991, S. 117, zitiert nach Busselle & Bilandzic, 2008).

Die Wahrnehmung und Interpretation der Berichte als authentisch und das Vertrauen, das Lernende diesen Berichten entgegenbringen, ist ein Erklärungsansatz für die besondere Wirkmächtigkeit von Zeitzeug:innenberichten (Bertram, 2017; Brüning, 2018; Heyl, 1995; Sabrow, 2012). Empirische Studien im Bereich der Medienpsychologie konnten Hinweise darauf finden, dass die Wahrnehmung der Medieninhalte als mehr oder wenig authentisch einen Einfluss auf die Verarbeitung der Inhalte hat: Konijn und Hoorn (2005) fanden Hinweise, dass authentische Medieninhalte, die auf Realität basieren, zu einem stärkeren Verbundenheitsgefühl mit den Protagonist:innen führten als fiktive Medieninhalte. Ji und Raney (2016) untersuchten den Zusammenhang zwischen wahrgenommener Authentizität, *suspension of disbelief* – der Bereitschaft, Zweifel auszublenden – und der Involviertheit mit den Medieninhalten. Die Ergebnisse ihrer Studie zeigten, dass die wahrgenommene Authentizität positiv mit *suspension of disbelief*, emotionalem und kognitiven Involvement zusammenhängt. Schließlich konnten Bilandzic und Busselle (2011) Hinweise darauf finden, dass die Zuschreibung von Authentizität Rezeptionserfahrungen in Form von Immersion, Identifikation mit den Charakteren und Eintauchen in die Geschichte fördert. Die Studien implizieren, dass die Wahrnehmung und Interpretation des Lernmaterials als authentisch einen Einfluss auf die Verarbeitungsprozesse der Lernenden haben.

Ich verstehe in der vorliegenden Arbeit die Authentizität des Lernmaterials im Kontext von Zeitzeug:innenberichten als eine Zuschreibung durch die Lernenden, von Menschen über Vergangenheit und Geschichte lernen zu können, die selbst dabei waren. Diese Zuschreibung hängt allerdings nicht nur von den Lernenden ab. Auch die Einführung und Inszenierung durch die Lehrkräfte sowie die Art und Weise des Auftretens der Zeitzeug:innen vor der Klasse – im Sinne der persönlichen Authentizität – sind bedeutsam. Mit der Zuschreibung des Berichts als authentisch geht ein besonderes Vertrauen einher, das Lernenden den Zeitzeug:innenberichten entgegenbringen und welches die Lernprozesse im Rahmen der Nutzung des Lernarrangements beeinflusst. Anknüpfend an empirische sowie konzeptionelle Überlegungen zur besonderen Wirkmächtigkeit von Zeitzeug:innenberichten nehme ich weiter an, dass die Zuschreibung der Authentizität eine kritische Auseinandersetzung mit den Erfahrungsberichten auf abstrakter Ebene im Sinne des historischen Denkprozesses erschwert. Jedoch bleibt dabei die Frage offen, inwiefern die Prozesse im Rahmen der Lernaktivität eine Rolle im Lernarrangement mit Zeitzeug:innenberichten spielen. Im folgenden Kapitel entwickle ich ein Modell, das verschiedene



Verarbeitungsmodi beim Lernen mit Zeitzeug:innenberichten identifiziert und im Kern die Lernaktivität im Lernarrangement mit Zeitzeug:innenberichten darstellt.

### **1.3.3 Lernaktivität: Vier Verarbeitungsmodi im Umgang mit Zeitzeug:innenberichten**

Die Lernaktivität umfasst sichtbare und mental ablaufende Lernprozesse (Seidel, 2014). Im Gegensatz zur Beschreibung des Unterrichtsangebots mit konkreten Merkmalen (Struktur-Paradigma; siehe Kapitel 1.1.2 und Kapitel 1.2.4), fußt die Idee der Nutzung des Angebots, also die Lernaktivität, auf dem Prozess-Paradigma. Hier geht es um Verarbeitungs- und Lernprozesse. Die Lernaktivität umfasst die Tätigkeiten der Lernenden im Unterricht. Chi und Wylie (2014) entwickelten eine Klassifikation von Lernaktivitäten als passive (z. B. zuhören), aktive (z. B. lesen und unterstreichen), konstruktive (z. B. Text produzieren) und interaktive (z. B. Argumentation in Interaktion) Tätigkeiten zur Förderung kognitiver Lernprozesse im Unterricht. Sie formulierten die Annahme, dass passive Aktivitäten den geringsten und interaktive Aktivitäten den höchsten Lernzuwachs im Sinne eines tiefgreifenden Verständnisses der Lerninhalte bringen. Über diese verhaltensbezogenen Aspekte der Lernaktivität hinaus sind auch die mental ablaufenden Verarbeitungsprozesse Teil der Lernaktivität. In diesem Zusammenhang unterschieden beispielsweise Fredricks, Blumenfeld und Paris (2004) zwischen verhaltensbezogenem, emotionalem und kognitivem Engagement der Lernenden im Unterricht. Aus ihrer Literatursynthese folgerten sie, dass sich verhaltensbezogenes Engagement in der Beteiligung der Lernenden an Unterrichtsaktivitäten, emotionales Engagement in affektiven Reaktionen der Lernenden auf verschiedene Stimuli (z. B. Verhalten der Lehrkraft, Inhalte) und kognitives Engagement vorwiegend in Strategien selbstregulierten Lernens äußern. Die Lernaktivität im Angebots-Nutzungs-Modell umfasst also beides: sichtbare Aktivitäten der Lernenden und mental ablaufende Verarbeitungsprozesse. Letzteres steht in der vorliegenden Dissertation im Fokus.

Denn unabhängig davon, ob die Lernenden den Bericht von Zeitzeug:innen lediglich hören und sehen oder das Interview mit Zeitzeug:innen in einer Projektarbeit möglicherweise selbst führen (Kapitel 1.2.2), die Informationen aus dem Bericht müssen von den Schüler:innen aktiv verarbeitet werden (Renkl, 2015): Sie müssen als relevant interpretiert, mit Vorwissen verknüpft und im Arbeitsgedächtnis verarbeitet werden, um ihren Weg ins Langzeitgedächtnis zu finden. Diese Verarbeitungsprozesse können darüber Aufschluss geben, wie es zu einer bestimmten Wirkung des Lernmaterials kommt und wieso Zeitzeug:innenberichte möglicherweise nicht zu einer Förderung historischer Kompetenzen beitragen (z. B. Bertram et al., 2017).

Zeitzeug:innenberichte sind Narrative – Geschichten, die etwas über eine Person und ihre tatsächlichen Erlebnisse erzählen, Orte beschreiben, zu einer historischen Zeit Bezug nehmen und dabei häufig einer narrativen Struktur folgen. Medienpsychologische Forschung bietet Anknüpfungspunkte für das Verständnis der Verarbeitungsprozesse beim Lernen mit Zeitzeug:innenberichten. In der Rezeption von narrativen Filmen können zwei Verarbeitungsmodi unterschieden werden: Im distanziert-analysierenden Verarbeitungsmodus denken die Rezipient:innen „über den Film“. Sie sind unbeteiligte Beobachter:innen. Im distanzlos-involvierten Verarbeitungsmodus denken die Betrachter:innen „in dem Film“. Sie sind fasziniert, emotional und kognitiv in das Geschehen involviert (Klimmt & Vorderer, 2003; Vorderer, 1992). Dieser involvierte Verarbeitungsmodus wurde in verschiedenen konzeptuellen Modellen weiter ausdifferenziert. Green und Brock (2000) beschrieben die Erfahrung, die durch ein solches Rezeptionsverhalten evoziert wird, als *transportation*, „a convergent process, where all mental systems and capacities become focused on events occurring in the narrative“ (Green & Brock, 2000, S. 701). Mit dieser Erfahrung gehen der Verlust an Wahrnehmung der Umgebung und eine starke emotionale Involviertheit in Bezug zu den Ereignissen in der Geschichte einher. Busselle und Bilandzic (2008) entwickelten das *Model of Narrative Comprehension and Engagement*, welches die Zusammenhänge zwischen den Verstehensprozessen und Rezeptionserfahrungen beschreibt. Sie argumentierten, dass Rezipient:innen in ihrer Beschäftigung mit dem Narrativ in die Welt der Geschichte eintauchen und diese nicht mehr als außenstehende Person, sondern von innen erleben. Dieser Perspektivenwechsel führe zu Erfahrungen wie *Flow* (Csikszentmihalyi, 1991) oder *Transportation* (Green & Brock, 2000).

Überträgt man diese Überlegungen auf die Begegnung mit Zeitzeug:innenberichten, können die Berichte die Lernenden entweder in die Geschichte transportieren, also einen distanzlos-involvierten Verarbeitungsmodus auslösen, oder Lernenden etwas *über* Vergangenheit und Geschichte vermitteln, also den distanziert-analysierenden Verarbeitungsmodus auslösen. Diese Verarbeitungsmodi decken allerdings noch nicht die Möglichkeit ab, persönlich involviert zu sein und gleichzeitig *über* die Geschichte zu lernen. Da diese Form der Verarbeitung für das Ziel von Geschichtsunterricht, durch die Beschäftigung mit Geschichte Orientierung in Gegenwart und Zukunft zu finden, am vielversprechendsten erscheint, ist deshalb im Kontext des Lernens mit Zeitzeug:innen eine Erweiterung notwendig.

Die aus der Sozialpsychologie stammende *Construal-Level-Theorie* (CLT; Liberman, Trope & Stephan, 2007; Mrkva, Travers & Van Boven, 2018; Trope & Liberman, 2010) beschreibt die wechselseitige Beziehung zwischen psychologischer Distanz (temporal, räumlich, sozial, hypothetisch) und Abstraktionsniveau mentaler Repräsentationen. Auf der Grundlage

dieser Theorie lässt sich ein Koordinatensystem erstellen, das die persönliche Distanz (persönliche Involviertheit) und das Abstraktionsniveau (meta-kognitive Involviertheit) in Beziehung zueinander setzt. Abbildung 3 zeigt das Modell: Die x-Achse beschreibt die wahrgenommene persönliche Distanz zur Vergangenheit. D. h.: Wie nah fühlen sich die Lernenden den Ereignissen, von denen die Zeitzeug:innen berichten, und wie emotional involviert sind sie in die Geschichte? Die y-Achse bildet das Abstraktionsniveau ab: Wie konkret bzw. abstrakt denken die Schüler:innen über die vergangenen Ereignisse aus den Berichten? In diesem Koordinatensystem lassen sich vier Verarbeitungsmodi verorten, die auch die beiden Verarbeitungsmodi nach Vorderer (1992) integrieren: ein distanziert-analytischer Modus, ein distanzlos-involvierter Modus, ein vergangenheitsbezogener Modus und ein gegenwartsbezogener Modus der Verarbeitung. Die Verarbeitungsmodi sind als dynamisch zu verstehen, d. h., Lernende verbleiben während und nach der Rezeption des Berichts nicht unbedingt bei einem Modus, sondern können durch verschiedene Stimuli (z. B. Inkonsistenzen im Bericht, Vergleich mit anderen Berichten, bestimmte Aufgabenstellungen durch die Lehrkraft) wechseln.

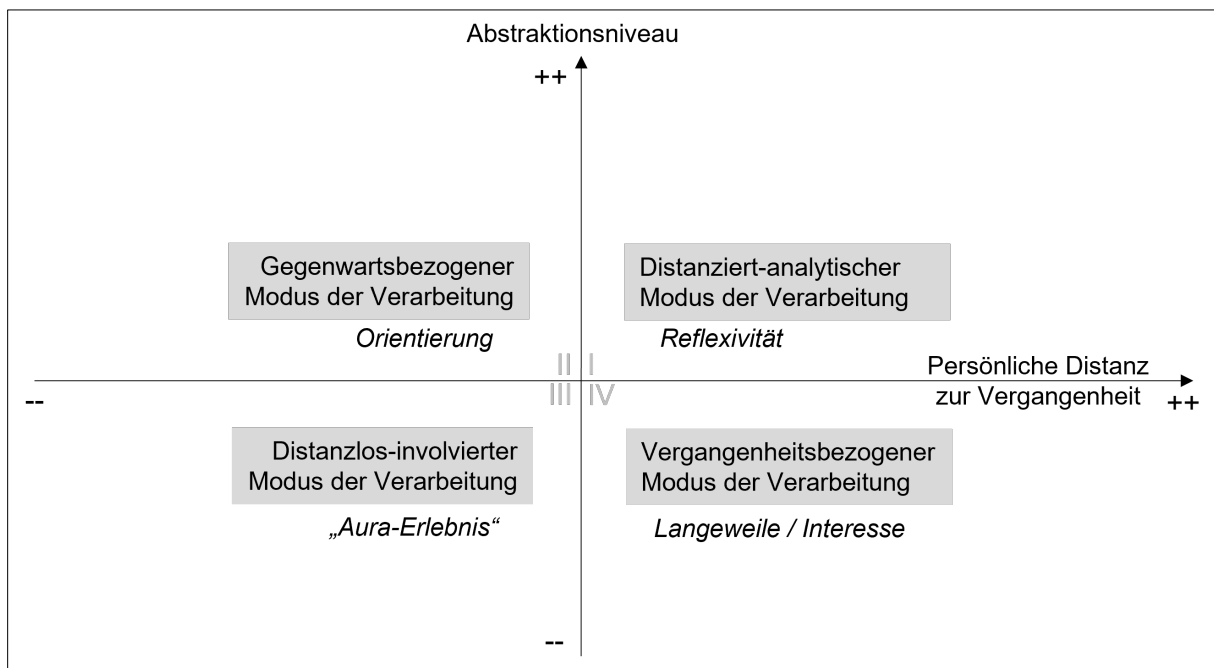


Abbildung 3. Vier Verarbeitungsmodi beim Lernen mit Zeitzeug:innenberichten

Empirische Befunde zu der Begegnung mit Zeitzeug:innenberichten lassen darauf schließen, dass viele Schüler:innen die Zeitzeug:innenberichte im distanzlos-involvierten Modus verarbeiten (III. Quadrant; siehe Kapitel 1.2.3; Ballis & Schwendemann, 2021; Bertram

et al., 2017; Brüning, 2018; Richardson, 2021). Das heißt, sie abstrahieren wenig und die Vergangenheit wird als sehr nah empfunden. So wie bereits von diversen Wissenschaftler:innen in der Medienpsychologie beschrieben, tauchen sie in die Geschichte ein, identifizieren sich möglicherweise mit den historischen Akteur:innen und denken „in der Geschichte“ (Vorderer, 1992; Green & Brock, 2000, Busselle & Bilandzic, 2008). In Anlehnung an Sabrows (2012) Überlegungen, nach denen Zeitzeug:innenberichte eine „Aura der Authentizität“ umgibt, kann diese Lernaktivität auch als „Aura-Erlebnis“ bezeichnet werden. Lernende sind so überwältigt von den individuellen Erfahrungsberichten, dass eine Kontextualisierung oder Dekonstruktion der Narrative ausbleibt (z. B. Ballis & Schwendemann, 2021; Bertram et al., 2017; Brüning, 2018; Richardson, 2021; siehe hierzu auch Kapitel 1.2.3). Verbleiben die Lernenden in diesem „Aura-Erlebnis“, besteht die Gefahr, dass sie möglicherweise Fehlkonzepte von Geschichte entwickeln.

Der distanziert-analytische Modus (I. Quadrant) zeichnet sich dadurch aus, dass die Schüler:innen eine Distanz zur Vergangenheit empfinden und die Informationen durch einen Vergleich mit weiteren Quellen abstrahieren. Dies geht häufig mit historischen Methoden- und Sachkompetenzen einher (z. B. Schreiber et al., 2006; siehe hierzu auch Kapitel 1.1.2). Vorderer beschrieb diesen Modus als distanziert-analysierende Verarbeitungsweise, in der Rezipient:innen *über* die Informationen nachdenken. Die Lernaktivität zeichnet sich in diesem Modus durch die Fähigkeit zur Reflexion aus. Allerdings haben Studien gezeigt, dass nur wenige Schüler:innen in der Lage sind, sich von dem Bericht zu distanzieren und die Informationen reflektiert mit ihrem Vorwissen und weiteren Berichten zu verknüpfen (Bertram et al., 2017; Brüning, 2018; Richardson, 2021).

Das Koordinatensystem bietet zwei weitere Verarbeitungsmodi, die für den Umgang mit Zeitzeug:innenberichten bedeutsam erscheinen. Im vergangenheitsbezogenen Modus der Verarbeitung (IV. Quadrant) nehmen Lernende die Informationen über die Vergangenheit in ihren Details auf, allerdings empfinden sie dadurch keine Nähe zur vergangenen Zeit oder der Person. Sie abstrahieren die Informationen aus den persönlichen Berichten nicht, sodass keine Bewertung der Informationen stattfindet. Hier ist denkbar, dass die Lernaktivität sich entweder durch Langeweile (im Sinne von: Geschichte interessiert mich nicht besonders, dennoch nehme ich die Informationen in ihren Details wahr) oder Interesse (im Sinne von: Ich bin sehr geschichtsinteressiert und mich interessiert jedes Detail aus der Vergangenheit) auszeichnet. Zahlen und Chronologien stehen hier im Vordergrund.

Schließlich ist anzunehmen, dass der Zeitzeug:innenbericht auch mit starkem Bezug zur Gegenwart verarbeitet wird: Das äußert sich im gegenwartsbezogenen Modus der Verarbeitung (II. Quadrant). Dieser Verarbeitungsmodus umfasst das, was Bildungspläne als zentrales Ziel von Geschichtsunterricht definieren: aus der Vergangenheit eine Orientierung in der Gegenwart finden (siehe hierzu auch Kapitel 1.1.2). Schüler:innen empfinden die vergangenen Ereignisse als sehr nah, erkennen die Relevanz der Geschichte für sich persönlich, abstrahieren sie gleichzeitig auf eine Ebene, die ihnen ermöglicht, sich selbst in der Gegenwart zu verorten. Sie bleiben folglich nicht bei den Details aus dem Bericht, sondern verknüpfen ihn mit aktuellen gesellschaftlichen Entwicklungen und der persönlichen Geschichte. Dieser Verarbeitungsmodus erfordert Orientierungskompetenz, und die Lernaktivität zeichnet sich durch ein Gefühl der Orientierung in Gegenwart und Zukunft aus.

Dem Verständnis dieser Verarbeitungsmodi liegt die Annahme zugrunde, dass durch die Gestaltung des Lernarrangements die verschiedenen Modi der Verarbeitung initiiert werden. Erfahrungsberichte über die Begegnung mit Zeitzeug:innenberichten sowie empirische Studien lassen allerdings darauf schließen, dass die besondere Wirkmächtigkeit von Zeitzeug:innenberichten in verschiedenen Formaten und zu verschiedenen Themen auf eine distanzlos-involvierte Verarbeitung der Berichte zurückgeht (Angvik & von Borries, 1997; Ballis & Schwendemann, 2021; Bertram et al., 2017; Richardson, 2021). Um zu verstehen, was diesen Verarbeitungsmodus ausmacht, bedarf es einer systematischen Konzeptualisierung und empirischen Erforschung. Nur so können die Lernprozesse in diesem Lernarrangement differenziert verstanden werden. Deshalb steht das „Aura-Erlebnis“ als Verarbeitungsmodus im Rahmen der Lernaktivität im Fokus der vorliegenden Dissertation.



---

# 2

ZIELSETZUNG  
UND  
FORSCHUNGSFRAGEN

---

## 2 Zielsetzung und Forschungsfragen

Eine Verknüpfung der generischen Perspektive der Empirischen Bildungsforschung auf Unterrichts- und Lernprozesse mit fachspezifischen Perspektiven der Geschichtsdidaktik verspricht einen hohen Erkenntnisgewinn für beide Disziplinen (Trautwein et al., 2021). In den vorherigen Kapiteln habe ich gezeigt, wie das generische Angebot-Nutzungs-Modell zur Beschreibung eines spezifischen Lernarrangements im Geschichtsunterricht dienen kann, um so ein umfangreiches Verständnis von Einflussfaktoren, Lernprozessen und Wirkungen zu erhalten. Das Lernarrangement mit Zeitzeug:innenberichten ist für Lernende wie Lehrende ein beliebter und eindrücklicher Zugang zur Geschichte, obwohl empirische Studien darauf hinweisen, dass dieser individuelle Zugang eher überwältigen kann, als ein reflektiertes Geschichtsverständnis zu fördern. Diese Ergebnisse werfen Fragen danach auf, was sich zwischen Unterrichtsangebot und Wirkung abspielt. Geschichtsdidaktische Forschung beschäftigte sich bisher allerdings kaum mit der empirischen Erforschung der Lernprozesse in dem Lernarrangement mit Zeitzeug:innenberichten: Die Lernprozesse erscheinen bisher vielmehr als Blackbox, als dass Wissenschaftler:innen wie Praktiker:innen sie genau verstehen. Die Perspektive der Empirischen Bildungsforschung kann zu der systematischen und empirischen Erforschung dieser Lernprozesse einen wichtigen Beitrag leisten. Zudem zeigen die vorherigen Kapitel, dass das Angebots-Nutzungs-Modell zur Beschreibung unterrichtlicher Wirkungsweisen wertvolle Anknüpfungspunkte für eine umfassende Erforschung des Lernarrangements und seiner Einflussfaktoren bieten kann. Das übergeordnete Ziel dieser Dissertation ist es, über die systematische Untersuchung des Lernarrangements ein tiefgreifendes Verständnis für die Lernprozesse beim Lernen mit Zeitzeug:innenberichten zu bekommen. Für dieses Ziel werden die Perspektiven von Empirischer Bildungsforschung und Geschichtsdidaktik verknüpft. Diese Verbindung bietet wertvolle Beiträge für beide Wissenschaftszweige.

Für die Erstellung der Dissertation waren folgende Forschungsfragen leitend, um die Blackbox der Lernprozesse zu betrachten: Welche Prozesse machen dieses Lernarrangement (a) im Bereich der Lernaktivität, (b) im Bereich der Wahrnehmung der Lernenden und (c) im Bereich des Zusammenspiels von Lernvoraussetzungen und Unterrichtsangebot aus? Insgesamt habe ich für die Beantwortung dieser Fragen zur systematischen Untersuchung des Lernarrangements und der zugrunde liegenden Prozesse vier Studien durchgeführt, die sich an verschiedenen Stellen des Angebot-Nutzungs-Modells einordnen lassen (für eine Übersicht siehe Abbildung 4): im Bereich der Lernaktivität (Studie 1, Studie 2, Studie 3), im Bereich der Wahrnehmung und Interpretation (Studie 2, Studie 3), im Bereich des Lernpotenzials (Studie 2,



Studie 4), im Bereich des Unterrichtsangebots (Studie 2, Studie 4) und schließlich im Bereich der Wirkungen (Studie 4).

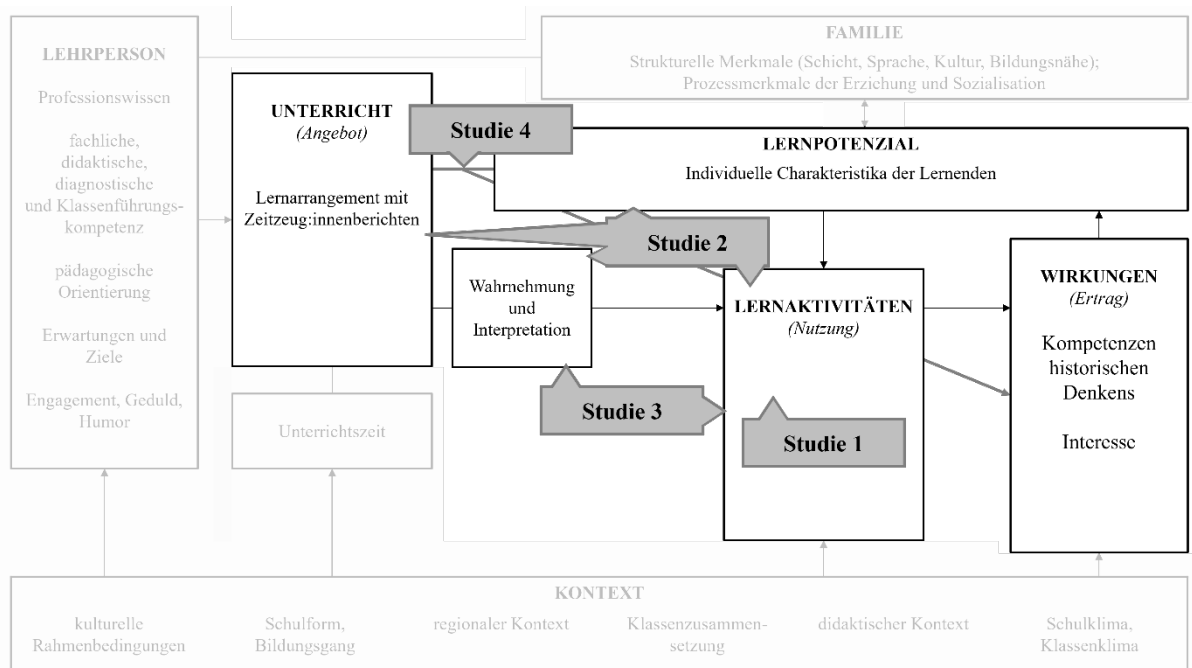


Abbildung 4. Einordnung der vier Dissertationsstudien in das Angebots-Nutzungs-Modell nach Helmke (2017)

**Studie 1** nimmt das Lernarrangement mit Zeitzeug:innenberichten aus theoretischer Perspektive in den Blick und folgt der Forschungsfrage, wie die Lernaktivität von Lernenden mit Zeitzeug:innenberichten strukturiert ist und welche Verarbeitungsprozesse die Lernaktivität auszeichnen. Die Studie befasst sich mit Zeitzeug:innenberichten und historischen Orten als komplexen historischen Quellen und entwirft ein konzeptuelles Modell, das die Lernaktivität mit kognitiver, affektiver und physischer Involviertheit beschreibt. Das konzeptuelle Modell basiert auf theoretischen und empirischen Studien der Geschichtsdidaktik, der (Medien-) Psychologie und der Anthropologie. In Studie 1 steht die Lernaktivität in Form des „Aura-Erlebnisses“ und ihre Beschaffenheit im Vordergrund. Sie ist die Grundlage für Studie 2 und Studie 3.

**Studie 2** baut auf dem konzeptuellen Modell auf und hat zum Ziel, ein Messinstrument zu entwickeln, das die Lernprozesse im Lernarrangement mit Zeitzeug:innenberichten standardisiert erfassen kann. Hier steht zunächst die Frage im Vordergrund, ob sich die theoretisch angenommenen Dimensionen (aus Studie 1) empirisch trennen lassen. Darüber hinaus beantworte ich in Studie 2 die Frage, in welchem Zusammenhang das „Aura-Erlebnis“ mit der distanziert-analytischen Verarbeitung, zu Lernvoraussetzungen und Merkmalen der Unterrichtsgestaltung steht. Datengrundlage hierfür bilden  $N = 668$  Schüler:innen der Klassen 8 bis

12, die nach der Begegnung mit einer:m Zeitzeug:in aus der ehemaligen DDR mit Hafterfahrung einen neuen Fragebogen zur standardisierten Erfassung der Lernaktivität ausfüllten. Mit dieser Studie entwickle ich ein standardisiertes Messinstrument und thematisiere den Zusammenhang der Lernaktivität mit individuellen Merkmalen der Lernenden (Geschlecht, aktive Mitarbeit) sowie situationalen Merkmalen der Unterrichtssituation (Intensität der Vorbereitung, Abstand zwischen Begegnung mit dem Zeitzeug:innenbericht und Erhebungssituation).

**Studie 3** thematisiert den Bereich der Wahrnehmung und Interpretation der Lernenden im Angebots-Nutzungs-Modell und geht der Frage nach, welchen Effekt ein zentrales Merkmal des Lernarrangements – die (wahrgenommene) Authentizität – auf die Lernaktivität der Lernenden hat. Diese Frage wird mit Daten einer experimentellen Online-Studie mit  $N = 452$  Teilnehmenden beantwortet, in der die Teilnehmenden jeweils einen Videoclip von einer:m Zeitzeug:in anschauen. Die Videoclips wurden den Teilnehmenden entweder als (1) originaler Bericht, (2) nachgespielter Bericht von Schauspieler:innen oder als (3) rekonstruierter Bericht durch Historiker:innen präsentiert. Die Videoclips zeigten jedoch alle originale Zeitzeug:innen. Studie 3 nutzt das theoretische Modell aus Studie 1 und den standardisierten Fragebogen aus Studie 2. Sie weist auf die Bedeutung der Wahrnehmung und Interpretation des Lernarrangements bzw. des Lernmaterials für die Lernaktivität hin.

In **Studie 4** steht schließlich die differenzielle Wirkung des Lernarrangements für Lernende mit unterschiedlichen Lernvoraussetzungen im Vordergrund. Konkret geht Studie 4 der Frage nach, ob sich die Wirkung des Formats von Zeitzeug:innenberichten für Lernende mit unterschiedlichen Lernvoraussetzungen unterscheidet. Diese Frage wird anhand einer randomisiert-kontrollierten Interventionsstudie mit  $N = 962$  Schüler:innen der 9. Klasse in Baden-Württemberg beantwortet (Bertram et al., 2017). Hier wird das Oberflächenmerkmal des Formats der Zeitzeug:innenberichte in Form von Videos, Texten oder einer Live-Begegnung untersucht. Als Lernvoraussetzung definiert die Studie das Wohlbefinden der Lernenden in sozialen Situationen, da sich durch die Formate unterschiedliche soziale Situationen ergeben. Die Wirkung wird durch drei Kompetenztests zum historischen Denken erfasst. Studie 4 verbindet im Angebots-Nutzungs-Modell das Unterrichtsangebot, die Lernvoraussetzungen und Wirkungen, berücksichtigt die Lernaktivität jedoch nicht. Dennoch trägt Studie 4 durch die differenzielle Betrachtung unterschiedlicher Schüler:innengruppen zu einem tiefergreifenden Verständnis der Wirkmechanismen im Lernarrangement mit Zeitzeug:innenberichten bei.

Zusammengenommen tragen die vier Studien durch die Verknüpfung von theoretischen Perspektiven verschiedener Disziplinen (Geschichtsdidaktik, Empirische Bildungsforschung,

---

Psychologie, Anthropologie), die unterschiedlichen methodischen Zugänge (theoretische Konzeption, Fragebogenkonstruktion, Experiment, Sekundärdatenanalysen) und die unterschiedlichen Fokusse im Rahmen des Angebots-Nutzungs-Modells zu einem systematischen und differenzierten Verständnis des Lernens mit Zeitzeug:innenberichten bei. Konkret steht dabei im Zentrum, welche Rolle Emotionen und Kognitionen in den Lernprozessen mit diesen authentischen Geschichten haben. Die vorliegende Dissertation zeigt zudem den Gewinn interdisziplinärer Perspektiven auf das geschichtsdidaktische Lernarrangement mit Zeitzeug:innenberichten auf.



---

# 3

## STUDIE 1

---

Zachrich, L., Weller, A., Baron, C. & Bertram, C. (2020).  
Historical experiences: A framework for encountering complex  
historical sources. *History Education Research Journal*, 17(2),  
243–275. <https://doi.org/10.18546/HERJ.17.2.08>

**Abstract**

We encounter information about the past in everyday life through movies, books, or complex historical sources—like historic sites or eyewitness accounts. Investigations of how visitors and learners engage with these complex historical sources have mainly focused on the “something special” of the encounter on the one hand and on the clear cognitive engagement on the other. Yet, we know little about what and how learners and visitors learn from these complex historical sources and the resultant historical experiences. However, it is an important precondition for further theoretical and empirical research to fully understand these experiences. The current article takes the first step in building an integrated model to understand from a situated embodied perspective the *historical experiences* derived from encounters with complex historical sources. Drawing on German and English-language literature across related disciplines, we conceptualized the experience within an interplay of cognitive, affective, and physical engagement. Within these dimensions, we identified responses that indicate the different elements of the historical experience and discuss limitations and avenues for further research.

Keywords: *Historical thinking, embodied learning, situated learning, historical consciousness, historic site, eyewitness of the past*

## Introduction

In daily life, we are inundated with information about the past. We encounter history in historical narratives in television documentaries, movies, books and/or games with historical topics, and through *complex historical sources*—like historic sites and/or eyewitness accounts of the past—that are comprised of multiple media and/or levels of interaction and engender intellectual, affective, and physical engagement in the learner. Yet, we know very little about what and how people learn with these complex historical sources. Investigations of how people engage with complex historical sources have traditionally sought to explain the encounters as life-changing peak experiences (e.g., Latham, 2013) or as atomized source work (e.g., Baron, 2012). While these elements are both essential for explaining parts of the experience of engaging with complex sources, more integrated approaches (Endacott and Brooks, 2013; Wineburg, 2001) are needed to truly “get at” the historical experience. Previous holistic models (Dierking and Falk, 1992; Hooper-Greenhill, 2004) offer generic outcomes for learning about museums, which are useful for thinking about the entirety of the experience in museums, but do not capture the complexity of the unique disciplinary experiences that unfold when encountering complex historical sources.

While differences exist in the particulars of how people engage with complex historical sources, their motivation for doing so is the same: People seek out historic sites and eyewitnesses because they regard them as the most “authentic” (Angvik and Borries, 1997; Jones, 2016; Rosenzweig and Thelen, 1998; Van Boxtel *et al.*, 2016) ways to connect directly to the people, objects, or places that stood witness to the past because “The bones are right there. The bones don’t lie” (Rosenzweig and Thelen, 1998: 106). These encounters are primary experiences unmediated by a video/computer screen or a text, rendering them distinct learning experiences (Greene *et al.*, 2015; Landström *et al.*, 2005; Ramlogan *et al.*, 2014; Reiß *et al.*, 2014). Attempts to discern and measure the effect of that sense of authenticity on learners and build a pedagogy that builds on it or deconstructs it, is complicated by the disparate verbiage around it, and the ineffability of that which researchers are trying to define. Struggling to define that sense of *something* special about an eyewitness account or a historic place, researchers have generated an array of terms with overlapping meanings: “aura of authenticity” (Sabrow, 2012, translated by the author), “the numinous” (Latham, 2013), “authentic” (Hampp and Schwan, 2014; Rössner and Uhl, 2012), “authority” (Rosenzweig and Thelen, 1998; Trofanenko, 2006) and so on.

Further, the personal and subjective nature of that sense of authenticity further complicate the notion. For example, how does the learner encountering the piece of The True Cross at the Jerusalem Chapel in Bruges, Belgium attribute its authenticity? Is it as a physical artifact that was present at the crucifixion of Christ or as a medieval relic around which believers over centuries have imbued greater meaning? Similarly, what is the proximity to historical events that an eyewitness must have in order to be considered an “authentic” source? Did they have to physically be removing pieces of the Berlin Wall when it came down or merely in East or West Berlin to be considered to have “been there”? Arguments can be made in either direction, and rest on a multitude of subjective factors. While there is a general sense in the research literature that there is something larger happening when people engage with these sources, it remains tantalizingly unknowable, with descriptions veering closer to poetry than science.

At the opposite pole sits the discussion of source work in history education, which has largely focused on exploring the use of historical documents from a cognitive perspective (e.g., Wineburg, 1991). While researchers are building on that work with other types and combinations of sources (e.g., buildings: Baron, 2012; Gussmann *et al.*, 2017; documents and images: Baron, 2016; paintings: Glaser and Schwan, 2015; sound: Lee *et al.*, 2015), the pace and line of this work poses considerable problems for understanding what people learn from historic sources. Although this is a fruitful and necessary path for understanding these isolated sources, outside of laboratory or formal schooling settings it is a rare occasion when individuals work with historical sources in isolation. Rather, the power of the experience of engaging with complex historical sources requires consideration of the full range of the embodied experiences—including the thoughts, emotions, and the sensory and physical engagement—learners employ to learn about the past. Therefore, we must consider how individuals engage with the complex historical sources, and what that tells us about the *historical experience* and how that shapes learners’ understanding of the past.

Although history education should enable students and educators to handle these sources competently, we know little about the actual effects that engaging with historic sites and eyewitnesses of the past have on learners besides some first hints on the particularity of these sources (Baron, 2013; Baron *et al.*, 2019; Bertram *et al.*, 2017; Savenije, 2016). Consequently, studies on the effectiveness of these immediate experiences with the complex sources on learning processes are needed. If we understand the components of these encounters better, we might be able to derive methods for educational practitioners about how to use these sources



---

effectively, minimize the risks that acritical oversimplification poses and thus, empower learners to become reflective citizens.

For the last thirty years, we have focused largely on understanding the cognitive underpinnings of historical sources. Specifically, this has meant using historical documents to assess individuals' historical thinking (Wineburg, 1991; 1998) or historical reasoning (Van Drie and Van Boxtel, 2008). While these and other researchers noted that there are affective elements of historical source work (e.g., Sakr *et al.*, 2016; Savenije and De Bruijn, 2017; Zembylas, 2016) that are both inherent in and complicating our understanding of what people learn, extant frameworks do not offer a systematic way to consider these experiences working with historical sources. The work of two of this paper's authors (Baron, 2012; Bertram *et al.*, 2017) has focused on understanding what people learn from complex historical sources—buildings, images, and eyewitnesses. With those contributions, we recognized the limitations of this line of inquiry about historical sources.

While these studies expanded the range of historical sources that researchers could consider, it became clear that even if we identified the cognitive underpinnings of how individuals work with every type of source found at an historic site or encounter with eyewitnesses, it would still not help us understand what people learn from the experience of being at a site or working with eyewitnesses. Rather, much of the power of these *historical experiences* requires consideration of the *situated embodiment* of the learner and the ways in which their experiences inform them about the past. Situated learning shifts the unit of analysis of learning from either the individual historic source or the learner towards an interaction between the individual and the source and/or environment in which it is set. These interactions encompass the cognitive, social, and cultural contexts of learning (Cobb and Bowers, 1999; Greeno and Engstrom, 2014; Hutchins, 1995; Kirk and Kinchin, 2003; Nardi, 1996; Resnick, 1987).

Similarly, embodiment is “grounded in the relationship between a system and its environment. The more [one] can perturb an environment and be perturbed by it, the more it is embodied” (Fong *et al.*, 2003: 149). Herein, embodiment is intended to highlight, in part, the frequently overlooked importance of the physical body and its role in the agency of the learner and the multi-sensory, multimodal interactions inherent in engaging with complex historical sources and associated historical experiences. While thought, affect, and bodily sensations have distinct elements to them, embodiment speaks to the interdependence and interactivity between them: an idea (thought) might make us angry (emotion) which we express on our faces and a

host of autonomic responses, such as quickening of the pulse or postural changes (physical; Maiese, 2014). Thus, while our framework attempts to delineate observable responses into cognitive/affect/physical responses individuals have when encountering complex historical sources, we recognize that within the body, these systems work in concert.

The historical experiences to which we refer are direct experiences with historical sources not mediated by a video screen or an interpretive text, such as a museum panel, that prompt a larger consideration of that sources' role, import or effect on historic events or persons: Standing on Omaha Beach, your feet sinking into the sand as the tide rolls in, pondering how men in heavy gear could charge up the beach; Listening to a Holocaust survivor describe what she felt when her camp was liberated; the brutal power of the noonday sun beating down on you, making visceral the inhumanity of slaves' laboring on a Louisiana rice plantation—these are not purely intellectual experiences leading to the development of mental models. These are complex sensory interactions between the learner and the presences of the eyewitness or physical environs of the historic place, contextualized by the learner's knowledge of the persons and events of the past. More than just cognitive engagement, affective experiences, or physical sensations, the complex interplay of all of these modalities play a considerable role in understanding complex historical sources and the historical experience. Yet, we currently have no model or mechanism for assessing historic sites or eyewitness encounters as sources at that level of complexity. In essence, we are missing a deep understanding of how people experience these encounters with the past, which is a crucial precondition for encouraging historical experiences that engage a deep learning process.

The conceptual model offered in this paper considers learners' cognitive, affective and physical engagement with what we have termed *complex historical sources*. Due to the subjective nature of determining authenticity, we rest our delineation of complex historical sources, such as historic sites and eyewitnesses of the past, as being as distinct from atomized historical sources, such as single documents, by their irreducible composite nature. Complex historical sources are comprised of multiple media, and/or levels of interaction, and they engender an interplay of intellectual, affective and physical engagement in learners. In this article, we use examples from our own research into how people learn from eyewitnesses and historic sites to illustrate the model, although the model is applicable for use with complex historical sources beyond those examples.

Considerable work has been done to understand how individuals read historical documents and images (e.g., Wineburg, 1991), and this work has been essential in shaping our

understanding of how people learn history. The framework offered here is intended to consider historical materials beyond traditional text. However, in certain circumstances, documents would merit consideration as complex historical sources. For example, the U.S. Declaration of Independence is a document that is housed in a grand rotunda in the U.S. National Archives Building in Washington D.C. In this instance, a visitor to the National Archives is more likely to encounter the document as an historical or cultural artifact, rather than as a text for analysis. Here, the Declaration of Independence becomes a part of the complex historical source that is the National Archives Rotunda, and the *historical experience* of engaging with those materials in that place.

Additionally, we use the term “engagement” to describe the range of possible ways in which someone could interact with a complex historical source, e.g., talking to an eyewitness, climbing the stairs of the Duomo, analyzing an artifact. Our model synthesizes research across multiple fields (including media studies, museology, anthropology, cultural geography, and educational psychology) in both English and German to help us frame learning encounters with complex sources presumed to have some assignation of historical authenticity that is satisfying to the learner. Unlike other models that focus on historical thinking or consciousness, which almost exclusively consider cognition, this model, written from a situated embodied perspective (Dawson, 2014; Korthagen, 2010; Wilson, 2002), reverses the microscope and allows us to consider the holistic historical experience when encountering complex historical sources.

### **Research Question**

With the current article, we aim at taking the first step to understand learning with complex historical sources by building a conceptual framework to describe and define the learning experience in detail, its characteristics and the processes involved. We were guided by the following research question: What elements comprise the experience of encountering complex historical sources? From a situated embodied perspective, we argue that engaging with complex historical sources involves the interplay of cognitive, affective and physical experiences.

### **Method**

To approach our research question, we conducted a comprehensive literature review, drawing upon a wide range of English- and German-language sources to identify and delineate the overlapping terms, their attributes, and limitations that are associated with the complex encounter. Next, we held a series of small conferences in the United States and Germany, inviting researchers and practitioners across the educational, historical, and psychological

landscape to discuss the range of questions that arose around learning with complex historical sources and experiences. The first conference, held in Germany, brought together approximately ten researchers from across a range of subjects, over the course of a week, to consider Walter Benjamin's notion of Aura (1965) and how the notion of the "authentic" drove engagement and learning in history. From this, a smaller subgroup gathered three more times in the United States and once more in Germany to consider how to capture learning related to authentic historical sources and experiences, and thus develop this framework. First, we conducted a systematic review of the research literature in history, social studies and museum education to identify fully articulated theories or frameworks related to what and how people learn history. To deepen our understanding of the educational research literature, we drew upon sources in philosophy, history, museology, medicine, and beyond. From there, we used snowball sampling to identify empirical and theoretical literature related to the ideas or phenomena that were emerging as essential. As part of the refinement process, at each stage of development, external reviewers, including individuals who participated in the initial conference, were asked to provide feedback to help shape the framework. Our queries to the reviewers related to the clarity, veracity, and coherence of the framework.

Finally, we synthesized the literature and discussions, identified crucial concepts from across different disciplines associated with the experience of interest and built an integrated model. The constructs we used are deduced and defined by the respective discipline. For simplicity, whenever possible we used extant terms (e.g., contextualization) that were already well-understood or had specific meaning in the field, rather than generating novel terms. Even within their respective disciplines, some of these constructs have ambiguous meanings, which made their use in our framework challenging. Our goal was to avoid overlapping constructs within our conceptual framework. Using respondents' accounts of their encounters with eyewitnesses and historic sites from existing data sets from our prior research, we tested and refined the model to ensure clarity in the categories.

The following sections give a brief overview on the key concepts and theories that inspired and guided the development of our framework. Drawing from the literature in media psychology, we adapted transportation theory (Green and Brock, 2000) and the model of narrative comprehension and engagement (Busselle and Bilandzic, 2008), both of which focus on the psychological processes and effects of narrative persuasion. We identified the aesthetic encounter with artworks, researched by Csikszentmihalyi and Robinson (1990) through an anthropological perspective, as a framework for a similar experience. Also coming from anthropology, specifically anthropological museum research, we use Cameron and Gatewood's

(2003) and Latham's (2007, 2013) work related to "numen-seeking", which identifies the near-spiritual connection that visitors seek at historic sites and with historic objects. We took these considerations on the very same topic (with regard to historic sites) as part of our framework, but we argue that we can add to this understanding through the consideration of cognition, emotion and embodied learning.

Another construct we explored to develop our framework is the concept of historical empathy (De Leur *et al.*, 2017; Endacott and Brooks, 2013, 2018; Huijgen *et al.*, 2017; Savenije and De Bruijn, 2017). Situated in social studies education, in particular history education in the United States, historical empathy describes the cognitive and emotional aspects of empathy, as distinct from purely human affective empathy, to foster learning about the past (Endacott and Brooks, 2013). Finally, we adapted the concept of presence, which describes a state of feeling to be in a mediated world without feeling the mediation. This concept comes from computer science and psychology, in particular virtual reality research (Frank, 2015; Lombard and Ditton, 1997). We argue that all of these different theories, as well as other constructs, contribute to a better understanding of the complex encounter with historic sites and eyewitnesses of the past, and thus, we have synthesized them into the conceptual framework that follows.

We assigned the identified constructs to three main dimensions: cognitive engagement, affective engagement and physical engagement. While situated embodied cognition would subsume affective and physical engagement into cognition, our model flattens those distinctions and positions cognition as a 'first among equals', so as to be able to identify observable behaviors or ideas across the range of possible responses.

We use the term "cognition" and "affect", rather than "thought" or "emotions", as these terms are drawn from the affiliated literature and indicate a recognition of the greater scope that cognition and affect encompass, compared to just singular thoughts or emotions, as well as the underlying interrelatedness of these dimensions. For example, according to Zembylas (2007), affect encompasses an emotional and a physical response to stimuli (i.e., crying at a sad film). Determining how to delineate the differences between when someone says "I'm sad" (emotional response) and then begins crying (physical response), versus someone starting to cry and then squeaking out an "I'm sad", or someone saying "I'm sad", but not crying, or crying, but saying nothing, becomes something of an infinite loop. In short, we acknowledge the interrelationship of these feelings and actions, and we offer ways to categorize observable behaviours without severing the underlying interconnections between them. Further, from this stance, being able to denote either a physical or an emotional response in the absence of the

other allows for the possibility of connection, but remains agnostic on whether those connections occurred in a particular instance.

Identified within these three dimensions are differentiated responses to the historical sources that appear frequently in people's interactions with them. We understand these responses not necessarily to occur in every experience but display a range and constellation of possible responses that might happen. We partially renamed these responses and provide our own definitions to satisfy the specific needs of our context of complex historical sources. Additionally, we provide model statements similar to those used by respondents to help delineate the type of response each category engendered, but also the way in which the participants positioned themselves in relation to the source—i.e., at arm's length or standing in the shoes of—as that distance was critical for understanding whether they were engaging with the source in ways that were predominantly cognitively, affectively, physically. Finally, we identify the unifying experience that draws together engagement across the three modalities. Although we tried to define the responses as distinct as possible, they are, of course, interconnected. We acknowledge that an individual's responses are embedded within personal and situational characteristics: different learners might respond differently to the complex source depending on their multiple subjectivities and experiences; also, different complex historical sources will presumably lead to different qualities of the historical experiences. The conceptual framework is displayed in Figure 1, summarized in Table 1 and will be explained in detail later in the article. Table 2 outlines the literature from which we draw our categories.

### **Framework for encountering complex historical sources**

#### **Cognitive engagement**

Cognitive engagement is an important part of the complex experience of visiting historic sites and engaging with eyewitnesses. Almost all of the concepts listed above, no matter which discipline they belong to, contain a cognitive part: transportation, for example, occurs when learners focus their mental capacities to develop mental models about the narrative (Busselle and Bilandzic, 2008; Green and Brock, 2000; Sweller, 2010). Csikszentmihalyi and Robinson (1990) describe the intellectual response to an artwork as a “cognitive rush” that leads to critical engagement with the details of the presented story. The concept of presence implies a judgement about, and cognitive involvement with, the mediated space (Lombard and Ditton, 1997). Lastly, performing historical empathy is a decidedly cognitive act, grounded in the processes of perspective recognition and contextualization (Endacott and Brooks, 2013). All of these concepts assume that the recipient cognitively processes the presented information.

---

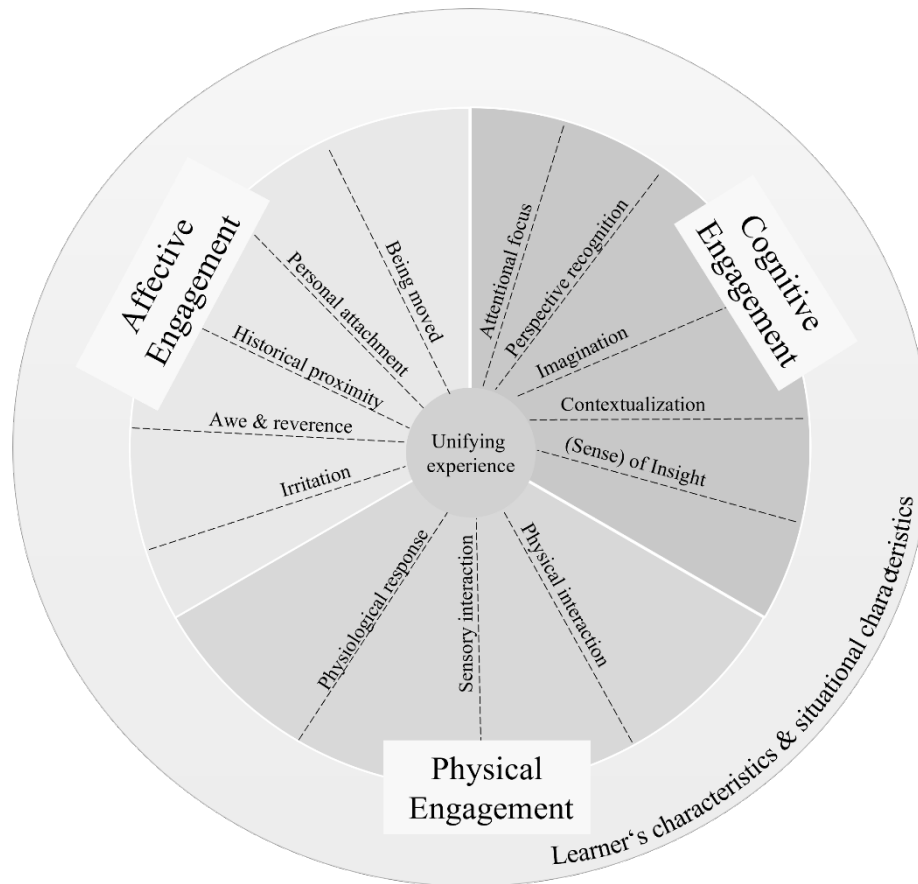
With regard to encountering complex historical sources, we assume that learners generate theoretical and historical questions about the lives of individuals and the historical and cultural context in which they were/are set, and begin to critically engage with the details of the site and story. Particularly, we argue that there are five distinguishable elements of cognitive engagement that are crucial and unique to the particular experience with complex historical sources: being deeply concentrated (e.g., Csikszentmihalyi and Robinson, 1990), imagining the time, place and people (e.g., Green and Brock, 2000), recognizing the perspectives of the historical agents (e.g., Endacott and Brooks, 2013), linking the information to one's prior knowledge (e.g., Huijgen *et al.*, 2017), and showing the insight to have learned something from the encounter (Busselle and Bilandzic, 2008). These responses are described in the following sections.

### *Attentional focus*

Attentional focus describes a state of deep concentration during an experience (Busselle and Bilandzic, 2008; Csikszentmihalyi and Robinson, 1990; Frank, 2015; Green and Brock, 2000; Lombard and Ditton, 1997). Learners report “losing themselves” in the experience, losing awareness of time and surroundings (Busselle and Bilandzic, 2008; Csikszentmihalyi and Robinson, 1990; Green and Brock, 2000; Latham, 2013; Lombard and Ditton, 1997). We understand attentional focus manifested in a deep concentration on the encounter with eyewitnesses and historic sites. Learners would describe this as ‘I fully concentrated on the object/story presented’ or ‘I did not even notice anything around me while interacting with the source’.

**Figure 1**

Conceptual framework of encountering complex historical sources



### *Imagination*

Imagination is a core prerequisite for narrative and historical understanding (e.g., Green and Brock, 2000; Klimmt and Vorderer, 2003; Lee, 1984; Lee and Ashby, 2001; Rösen, 2005). Referring to Schörken (1998), Brauer (2016a: 37) posits imagination as a mental capacity which “plays a role in every act of interpreting, receiving and reconstructing the past”. It is a form of visualizing something or somebody to make something distant more familiar in order for a better understanding. Imagination is not a flight of fancy, but a crucial part of the ability to construct the world of a person in order to understand their circumstances (Brauer, 2016a, 2016b; Csikszentmihalyi and Robinson, 1990; Klimmt and Vorderer, 2003).

We understand imagination as creating a mental imagery of the past and conjuring up a scene (Busselle and Bilandzic, 2008). Learners simply had a clear mental image of the place, story or time the site or historical agent witnessed. This can be understood as ‘I can imagine the time, the eyewitness talked about’, ‘I could picture how this place must have been looked like’ or ‘I had some clear pictures about the scene of the past in my head’.



**Table 1**

Summary of the relevant elements of the framework of engaging authenticity: Dimension, response, definition and example statement

<b>Dimension</b>	<b>Response</b>	<b>Definition</b>	<b>Typifying statement</b>
Cognitive engagement	Attentional focus	Learners deeply concentrate on the encounter with the source.	“I did not even notice anything around me while interacting with the source.”
	Imagination	Learners create a mental image or envision a particular version of the information presented by the source.	“I could imagine exactly, how the place looked like 100 years ago”, or “I could picture the scene the eyewitness talked about.”
	Perspective recognition	Learners understand the view of the historical agent (the eyewitness or a person who was involved with the site).	“I understand why the historical agent who spent time at this place felt and acted that way”, or “I totally understood the eyewitness’s thoughts and feelings.”
	Contextualization	Learners link the displayed information to their prior knowledge about the historical time/place/events.	“When I think of the circumstances the people had to deal with during that time, their actions become even more courageous...”
	(Sense of) insight	Learners believe they have a better understanding about the past because of new information/perception gained from experience with the source.	“I know better now how life after war must have been like.”

<b>Dimension</b>	<b>Response</b>	<b>Definition</b>	<b>Typifying statement</b>
Affective engagement	Being moved	Learners are emotionally touched by the information the source conveys.	“The object in the exhibition moved me”, or “The story of the eyewitness deeply touched me.”
	Personal attachment	Learners identify with the historical agent through a personal connection or perceiving the story from their point of view.	“The historical agent’s story reminded me of my grandmother”, or “I could feel the feelings of the eyewitness myself.”
	Awe and reverence	Learners <i>feel</i> a deep appreciation for the historical agent and a connection bigger than themselves.	“My God, when you think of what they went through.”
	Historical proximity	Learners feel a spatial and temporal closeness to the past.	“The past became vivid.”
	Irritation	Learners have a feeling of irritation because of unexpected or conflicting information.	“I wasn’t expecting that”, or “I always thought it has been different.”
Physical engagement	Physiological response	Learners react to the complex sources with involuntary bodily responses.	“Being at the site gave me an uncomfortable feeling in my stomach”, or “The story gave me goosebumps.”
	Physical interaction	Learners physically move through the site and/or have a conversation with the eyewitness.	“I walked the same way xyz did”, or “I could get answers to my very personal questions.”
	Sensory interaction	Learners perceive the elements of the complex sources within an interplay of their senses.	“I could smell the eyewitness’s perfume”, or “The museum felt cold and dark.”
Uniting	Unifying experience	Learners perceive the experience as a unifying moment of cognitive, affective and physical engagement.	“I had a moment where just everything came to me and taught me more than I could have learned with hundred books.”

---

*Perspective recognition*

More specific than imagination is the response of recognizing the perspectives of the historical agents. Scholars across different theoretical positions contend that historical perspective taking—“the attempt to understand the motives, beliefs, and behaviors of people in the past” (Kohlmeier, 2006; 34)—is a crucial component for understanding the past (Barton and Levstik, 2004; Cameron and Gatewood, 2003; Endacott, 2010; Endacott and Brooks, 2013; Hartmann and Hasselhorn, 2008; Huijgen *et al.*, 2017; Lee and Ashby, 2001).

Drawing from these conceptualizations, we define perspective taking as recognizing the perspectives of the historical agent, and understanding their thoughts, actions and decisions. Learners would describe this as ‘I see their point of view’ or ‘I can understand why the historical agent responded to the situation the way they did’.

*Contextualization*

At its most elemental, contextualization is the ability to situate events and agents in the full complexity of their historical time and place (Endacott and Brooks, 2013; Huijgen *et al.*, 2017; Van Drie and van Boxtel, 2008; Wineburg, 1991, 2001). Enacting contextualization requires learners to “[bring] forward multiple elements of prior knowledge of a particular time period—political positions, social conventions, economic forces, cultural and linguistic traditions—to understand the particular circumstances of the time and place” (Baron, 2016: 516).

We understand contextualization as the linkage of the information displayed at the site or by the eyewitness to one’s prior knowledge. We argue that through the linkage of information to prior knowledge of historical time periods and circumstances contextualization allows for the learner to fully experience these complex historical sources. Statements describing this are, for example: ‘I understand the historic agent’s circumstances’, ‘I understand the circumstances and constraints on the historic agent(s) because of the time/place they lived in’.

*(Sense of) insight*

Recipients who experienced transportation into a narrative return were somehow changed by the experience (Green, 2004). The experience with historic sites, as described by Latham involves “realizations about oneself, one’s identity, and one’s purpose in life” (Latham, 2013: 10). The aesthetic encounter with an artwork suggests that recipients learn something from the encounter, as well as experiencing transportation (Csikszentmihalyi and Robinson, 1990; Green, 2004). Additionally, visitors report to seek—among other things—learning

experiences at historic sites (Cameron and Gatewood, 2003), and students report to have learned a lot when encountering eyewitnesses of the past (Bertram *et al.*, 2017; Dutt-Doner *et al.*, 2016). Baron (2012) identified the moment of cognitive empathetic insight as the response to a physical stimulus provided by the historic site, indicating a malleability in how the experience occurs and is processed.

The encounter with a complex historic source leads the viewer to some kind of interpretative insight or profound understanding, leading to new and unexpected ideas in which “the individual apprehends something ordinarily beyond his or her capacities” (Thrash and Elliot, 2004: 957; see also Csikszentmihalyi and Robinson, 1990; Muth and Carbon, 2013; Perry, 2002; Schindler *et al.*, 2017). Referred to as the “aesthetic aha effect” (Muth and Carbon, 2013), people experience spontaneous inspiration and insight, enhancing their processing of information.

The (sense of) insight, we argue, can be part of the learning experience with complex historical sources that might occur as a moment of understanding which is a result of the intellectual interaction with the content displayed. This can occur in the form of an understanding about the distant past, one’s potentially changing conception of the world or oneself in life. Learners would describe it like ‘I suddenly knew how life must have been like’, ‘I have a better understanding of the past now’, or ‘I gained a better understanding about where I am in life’. An even deeper dimension of the understanding and insight as part of this complex experience is a *change of perspective*. This might be expressed by ‘It changed my perspective on life’ or ‘I have a totally different understanding of the past now’.

**Table 2**

Summary of literature

<b>Foundational concepts</b>		
<b>Concept</b>	<b>Cited literature</b>	<b>Empirical or theoretical</b>
Transportation theory	Green and Brock (2000)	Empirical
Narrative comprehension and engagement	Busselle and Bilandzic (2008)	Theoretical
Aesthetic encounter with artwork	Csikszentmihalyi and Robinson (1990)	Empirical
Numen-seeking	Cameron and Gatewood (2000)	Empirical
	Latham (2007)	Theoretical
	Latham (2013)	Empirical
Historical empathy	De Leur <i>et al.</i> (2017)	Empirical
	Endacott and Brooks (2013)	Theoretical
	Huijgen <i>et al.</i> (2017)	Empirical
	Savenije and De Bruijn (2017)	Empirical
Presence	Frank (2015)	Empirical
	Lombard and Ditton (1997)	Theoretical
<b>Elements of the framework</b>		
<b>Cognitive engagement</b>	<b>Cited literature</b>	<b>Empirical or theoretical</b>
Attentional focus	Busselle and Bilandzic (2008)	Theoretical
	Csikszentmihalyi and Robinson (1990)	Empirical
	Frank (2015)	Empirical
	Green and Brock (2000)	Empirical
	Lombard and Ditton (1997)	Theoretical
Imagination	Brauer (2016a)	Theoretical
	Brauer (2016b)	Theoretical
	Busselle and Bilandzic (2008)	Theoretical
	Green and Brock (2000)	Empirical
	Klimmt and Vorderer (2003)	Theoretical
	Lee and Ashby (2001)	Empirical
	Rüsen (2005)	Theoretical
Perspective recognition	Barton and Levstik (2004)	Theoretical
	Cameron and Gatewood (2003)	Empirical
	Endacott (2010)	Empirical
	Endacott and Brooks (2013)	Theoretical
	Huijgen <i>et al.</i> (2017)	Empirical

	Kohlmeier (2006)	Empirical
	Lee and Ashby (2001)	Empirical
Contextualization	Baron (2016)	Empirical
	Endacott and Brooks (2013)	Theoretical
	Huijgen <i>et al.</i> (2017)	Empirical
	Van Drie and Van Boxtel (2008)	Theoretical
	Wineburg (1998)	Empirical
	Wineburg (2001)	Empirical
(Sense of) insight	Csikszentmihalyi and Robinson (1990)	Empirical
	Green (2004)	Empirical
	Latham (2013)	Empirical
<b>Affective engagement</b>	<b>Cited literature</b>	<b>Empirical or theoretical</b>
Being moved	Brauer and Lücke (2013)	Theoretical
	Cameron and Gatewood (2013)	Empirical
	Dutt-Doner <i>et al.</i> (2016)	Empirical
	Latham (2007)	Theoretical
	Latham (2013)	Empirical
	Krämer (2012)	Theoretical
	Menninghaus <i>et al.</i> (2015)	Empirical
	Toila-Kelly (2018)	Theoretical
	Uhl (2012)	Empirical
Personal attachment	Brauer (2016a)	Theoretical
	Bilandzic and Busselle (2011)	Theoretical
	Cameron and Gatewood (2003)	Empirical
	Cohen (2001)	Theoretical
	Endacott and Brooks (2013)	Theoretical
	Mason <i>et al.</i> (2018)	Empirical
	Rosenzweig and Thelen (1998)	Empirical
Awe and reverence	Cameron and Gatewood (2003)	Empirical
	Gatewood and Cameron (2004)	Empirical
	Jones (2016)	Empirical
	Keltner and Haidt (2003)	Empirical
	Latham (2007)	Theoretical
	Latham (2013)	Empirical
	Sabrow (2012)	Theoretical
Irritation	Vogl <i>et al.</i> (2019)	Empirical
	Rose (2016)	Theoretical

Historical proximity	Benjamin (1965)	Theoretical
	Grever (2018)	Empirical
	Lombard and Ditton (1997)	Theoretical
	Klimmt and Vorderer (2003)	Theoretical
	Savenije and De Bruijn (2017)	Empirical
<b>Physical engagement</b>	<b>Cited literature</b>	<b>Empirical or theoretical</b>
Physiological response	Latham (2013)	Empirical
	De Manzano <i>et al.</i> (2010)	Empirical
	Nacke and Lindley (2008)	Empirical
	Waterton (2018)	Empirical
Physical interaction	Kunter and Trautwein (2013)	Theoretical
Sensory interaction	Arnold-de Simine (2012)	Theoretical
	Drozdzewski <i>et al.</i> (2016)	Theoretical
	Latham (2013)	Empirical
	Violi (2012)	Theoretical
	Waterton (2018)	Empirical

### Affective engagement

Affective engagement with the past comprises a significant element of learners' encounters with historic sites and eyewitnesses (Boler, 1999; Cameron and Gatewood, 2003; Gatewood and Cameron, 2004; Green and Brock, 2000; Latham, 2007, 2013; Mason *et al.*, 2018; Savenije and De Bruijn, 2017; Wetherell *et al.*, 2018; Zembylas, 2016, 2018). These sources are able to create an emotional involvement (von Plato, 2009) and have the potential to "not only reach the head of the students but also the heart" (Uhl, 2012: 279, translated by the author). Affective engagement with the past has been shown to engender "highly connective experiences" between the learner and the historic site or eyewitness (Latham, 2007, 2013). In our framework, we identified five distinct responses making up the affective engagement with the sources: a feeling of being moved (Menninghaus *et al.*, 2015), a feeling of personal attachment to the historical agent (Brauer, 2016a; Cohen, 2001; Endacott and Brooks, 2013), a feeling of historical proximity (Benjamin, 1965; Lombard and Ditton, 1997), a feeling of awe and reverence (Cameron and Gatewood, 2003; Jones, 2016; Keltner and Haidt, 2003; Latham, 2013), and a feeling of irritation (Rose, 2016; Vogl *et al.*, 2019). It is important to note that not all affective engagement is positive. Terrible things can inspire awe. One can be moved to anger as well as empathy and joy.

*Being moved*

Historical pedagogy accredits historic sites and eyewitnesses of the past the chance to reach people through an emotional channel (e.g., Brauer and Lücke, 2013; Krämer, 2012; Uhl, 2012; Toila-Kelly, 2018). Independent of the content of the historical source, learners report a moving experience after talking to an eyewitness or visiting a historic site (Cameron and Gatewood, 2003; Dutt-Doner *et al.*, 2016; Latham, 2007, 2013). Drawing on our preliminary analysis, we developed the concept of being moved to better grasp how this experience occurs.

We argue that the affective engagement with these complex sources engenders the feeling of being moved (Menninghaus *et al.*, 2015) by the story or site. Being moved refers to an emotional arousal in general, independent of emotional valence such as positive or negative feelings. Rather, this presents a delineation allowing for the full range of possible emotions evoked or observed. Learners would most likely respond to the information with statements such as ‘I am touched by the story’ or ‘The story of the historical agent deeply moved me,’ as well as statements such as ‘I was moved standing on the same spot where history took place’.

*Personal attachment*

A personal connection to the past has been found to be a driving force in why people seek out historic sites (Cameron and Gatewood, 2003; Endacott and Brooks, 2013; Mason *et al.*, 2018; Rosenzweig and Thelen, 1998). Feeling connected to a historical agent allows the learner to transcend the present and make the distant past more comprehensible, encouraging the learner to consider the full humanity of historic agents, and leading to a deeper understanding than cognitive engagement alone (Cameron and Gatewood, 2003; Endacott and Brooks, 2013). Bilandzic and Busselle (2011) refer to this connection as identification, a process of simulation and adoption of an historical agent’s goals and the experience of the emotions. Cohen (2001: 251) defined this phenomenological process as “a process that consists of increasing loss of self-awareness and its temporary replacement with heightened emotional and cognitive connections with a character”.

We argue that personal attachment can appear in the form of affective connection as referred to by Endacott and Brooks (2013), Rosenzweig and Thelen (1998) and Brauer (2016a): a specific element in a person’s life—a story, experience, hobby, appearance, geographic proximity—that is similar to the historical agent’s. It can also appear in the form of losing one’s own identity by “temporarily substituting one’s own perspective with another person’s perspective on events, people, and emotions in that other world” (Bilandzic and Busselle, 2011: 34; de Leur *et al.*, 2017). At its most extreme, the latter might lead to an over-identification or



---

uncritical identification that directs the learners to make faulty moral or ethical judgments about the past (Brauer, 2016a, 2016b; Van Nieuwenhuysse, 2011).

We understand this affective response as a feeling of personal attachment of the learner to the historical event or historical agent because of personal experiences or connections to his or her own life. This could be stated as ‘I feel a personal connection to the story presented’, ‘Our stories (mine and the historical agent’s) are very similar’, ‘The eyewitness reminds me of a loved one.’ However, it could go as far as ‘I could feel the feelings of the historical agent myself’.

This response is closely and dynamically tied to *perspective recognition* and *contextualization* (see above). Taken together, these elements can be understood to describe the concept of empathy as defined by Brauer (2016a) as well as Endacott and Brooks (2013): the personal attachment leads the learner to “de-distance” him/herself from the historical agent—an immediate and affective response; whereas perspective recognition and contextualization guide the learner to “distance” him/herself from the historical agent—a cognitive and reflective process.

#### *Historical proximity*

The concept of presence as used in virtual reality research refers to the feeling of being in a mediated space—without recognizing the medium (Lombard and Ditton, 1998): “An illusion of a non-mediated spatial environment or social entity” (Klimmt and Vorderer, 2003: 349). We adapt the concept of presence and argue that the medium is not responsible for the perception of a physically real surrounding through the medium (Frank, 2015), but for a strong feeling of temporal and spatial proximity to the place and time (Grever, 2018). This feeling of proximity is also connected to the idea of “aura” as originally defined by Walter Benjamin (1965) with regard to an original piece of an artwork. Benjamin ruminated: “What is an aura? A strange web of space and time: the unique appearance of a distance, as close as it may be” (57, translated by the author). In line with construal-level theory (Trope and Liberman, 2010), the sense of distant past is malleable based on contextual factors and degrees of abstractness that the learner experiences (Trope and Liberman, 2010; Van Boven *et al.*, 2010).

We argue that the nature of the complex sources, their embodiment of the past, makes the learner feel engaged through what Savenije and de Bruijn (2017: 3) call “mnemonic bridging”, linking the past with the present. We suggest that this spatial and temporal proximity is understood in statements such as ‘I could feel the distant past’ or ‘The past became vivid’.

*Awe and reverence*

Visitors of historic sites report “a feeling of smallness or a sudden understanding of the grandeur in the meaning of the world around them” (Latham, 2013: 14). Awe and reverence refer to learners’ sense of “being in the presence of something holy” (Cameron and Gatewood, 2003: 67f.), having a ‘spiritual communion’ (Gatewood and Cameron, 2004: 208), or being ‘carried away’ (Latham, 2007: 257). In relation to a historic place it is a “transcendental experience” that people can have in contact with a historic site or objects in an exhibit (Cameron and Gatewood, 2003: 110). Learners experience the “transcendental reaction” of being “struck by the power” of a place or object (Gatewood and Cameron, 2004: 211). In encountering eyewitness testimony, Sabrow (2012: 27, translated by the author) describes a similar experience as the power of an “aura of authenticity”; the more devastating their story, the more overwhelming their appearance for the listener.

We conceptualize awe and reverence as a deep, numinous appreciation for historical agents and their circumstances or a sense of larger meaning of associated with those persons, places or events. It refers to a sudden feeling about the appreciation for historical agents, one’s fundamental feelings of oneself and about the world. Here, learners will often exclaim their wonderment in terms that grasp at the meaning of what they are experiencing: ‘My God, when you think of what they went through’, ‘Just to stand in this same place...’.

*Irritation*

The previous affective engagements could be experienced as having a positive valance. Yet, in certain situations, the information encountered either in an exhibit or by an eyewitness of the past can present the learner with new information that may challenge and contradict their prior knowledge. Bringing about an element of surprise, this unanticipated or contradicting information might lead to cognitive incongruity (Vogl *et al.*, 2019). For some, the cognitive incongruity they are confronted with when encountering historic sites or eyewitnesses of the past results in irritation.

Irritation as an affective response arises when the learner has an alternate understanding or different attitudes about the past than they are encountering with a historic site or the eyewitness. Moreover, irritation could either result in further curiosity and effort to resolve the incongruities and building up of new knowledge (Vogl *et al.*, 2019), or lead to resistance where one ignores the new information and clings to their prior beliefs as an act of defiance (Rose, 2016). In essence, we argue that a possible response to the encounter with eyewitnesses and historic sites is the *feeling of irritation* due to cognitive incongruities between their prior

knowledge or expectations and the new information. Learners would most likely state this response like: ‘I expected something different’ or ‘I was confused by the new information’.

### Physical Engagement

The experience of engaging with historic sites and eyewitnesses are inextricably tied to the body through sensory interaction, physical interaction and physiological arousal (Busselle and Bilandzic, 2008; Cameron and Gatewood, 2003; Ellsworth, 2005; Joy and Sherry, 2003; Latham, 2013; Manzano *et al.*, 2010; Nacke and Lindley, 2008; Rosenberg, 2007). For example, physically navigating through a historic site—walking, engaging with hands-on displays—“inscribes the body in place and how our relationship to place, in turn instigates a particular kind of remembering grounded in the physical space of our present situation” (Rosenberg, 2007: 54). Similarly, interactions with eyewitnesses engage the senses, engendering visceral, “flow”-like sensations when hearing their stories (Joy and Sherry, 2003; Latham, 2013; Manzano *et al.*, 2010; Nacke and Lindley, 2008). We argue that learners who engage with complex historical sources physically engage with them through all their senses to respond to and interact with historical objects, places, and people. In particular, we distinguish between sensory interaction, physical interaction, and physiological responses.

### *Sensory interaction*

Sensory interaction involves a component of embodied learning that takes place in the immediate encounter of complex historical sources. We draw from Latham’s (2013) conceptualization of the numinous experience people feel when they encounter historic sites to further clarify sensory interaction with authentic historical sources. Latham describes a wide range of sensory engagements, including “visually and spatially perceived elements... “the need to soak it up with one’s eyes”” (Latham, 2013: 10). Sensory interaction can be found in the “intermediality” or dynamic that develops between the historic site/eyewitness and the learner’s experience (Arnold-de Simine, 2010). With regard to eyewitnesses of the past, the sensory interaction might express itself by being able to fully observe the eyewitness while talking, and his/her reaction to one’s question.

Drawing on the work of Drozdowski *et al.* (2016: 447) we distinguish sensory interaction with historical sources in that they can “be smelt, touched, felt, imagined, tasted, and heard”. We define sensory interaction as the interplay between the body’s senses such as smell, taste, touch, and sound with a historical source, such as a historic site or eyewitness. Sensory interaction can be found in phrases such as ‘I was struck by the smell of the shoes’ or ‘the museum felt cold and dark’.

*Physiological response*

Distinct from sensory interaction, a physiological response can be understood as an immediate, involuntary bodily reaction that one experiences when encountering complex historical sources. These encounters frequently engender some form of a physiological response including “having a rush, a feeling of blood to the face, butterflies, tingly excitement, being overwhelmed...” (Latham, 2013: 15) all the way up to a form of flow, or optimal experience, which includes attentional, existential, and temporal dimensions (Csikszentmihalyi and Robinson, 1990).

In addition to learners’ own descriptions of these physiological states, these embodied responses are observable, and possibly measurable with instruments like EEGs or ECGs (Manzano *et al.*, 2010; Nacke and Lindley, 2008). Evidence of physical arousal would be described by phrases like ‘I teared up’, ‘it made my heart beat fast’, ‘it gave me a pit in my stomach’, or ‘it gave me goosebumps.’ It is important to note, however, that these physiological responses are connected to the affective engagement to the complex historical sources. For example, the feeling of goosebumps occurs because the depth of a story moves us; this is different than experiencing goosebumps when walking into a room because it is cold.

*Physical interaction*

Distinct from these two responses are the active physical interactions that comprise the significant element of the embodiment of the historical experience. Learning does not only depend on the offer but on how the offer—of encountering complex historical sources—is used by the learner (Kunter and Trautwein, 2013). Navigating one’s way through the space, feeling the need to touch the objects, receiving impulses from the site and following them are physical interactions that are often carried out by the learners/visitors in order to affect their understanding of the site. In terms of eyewitnesses, this physical interaction could present itself as asking questions and having a conversation with the eyewitness. How these interactions with complex historical sources are used by the learner is a crucial part of the learning experience. Statements indicating this are for example: ‘I could get answers to my own questions’, ‘I walked the same paths the soldiers walked at Omaha Beach’.

## Framing dimension: Unifying experience

“The information in the work of art fuses with information in the viewer’s memory—followed by the expansion of the viewer’s consciousness, and the attendant emotional consequences” (Csikszentmihalyi and Robinson, 1990: 38). The aesthetic encounter is a multidimensional experience that “integrates the visual with the emotional and the intellectual”

(Csikszentmihalyi and Robinson, 1990: 83). Latham (2013) called one dimension of the numinous experience “unity of the moment”, which is the holistic experience that frames the other responses as part of the experience. This moment involves intellect, experiences, emotions, and physical responses. It is deep, dynamic and vivid (Latham, 2013: 9): “The uniting experience is not a connection flowing through the experience, it *is* the experience”. Latham described it further as “one whole swirling entity of these things, overlapping and connecting. It is the uniting of all these things – emotion, intellect, feeling, senses, imagination—that results in meaning for the experiencer” (11). This is the intertwining of all facets of the experience responses, resulting in a sense of wholeness. It might even result in an inability to articulate the distinguishing elements of the experiences.

Based on these considerations, we argue that the overall interaction between cognitive, affective and physical engagement of the learner results in a unifying experience. Although hard to articulate, learners most closely would describe this with phrases like ‘It happened all at once’ or ‘I had a moment...’.

### **Application**

In order to show how this framework could help researchers, we offer two examples of how to use the framework to code learner responses. These examples are drawn from our existing research data sets and were chosen for the variety of framework elements they represent and compactness (i.e., the greatest range of responses in the shortest section of text). Note that we indicate ways in which coding can overlap to draw out the depth of the experiences in which the learners engaged.

The first example (see Table 3) is a female U.S. high-school teacher on a study-abroad professional development tour reflecting on her visit to the Hiroshima Peace Museum in Japan, a site that explores the effects of the atomic bomb dropped on Hiroshima during the Second World War by the American armed forces. This teacher has 18 years of classroom experience and a master’s degree in American history. We intentionally chose this example to show what a strong emotional response to an historic site might be, even for someone who has deep prior knowledge of the site and the events relating to it.

The second (see Table 4) is an interview with a 15-year-old female student who participated in an in-school intervention with eyewitnesses of the past talking about their roles in the opposition movement in the Peaceful Revolution of the German Democratic Republic. The interview captures the student’s impression of the encounter and her feelings about meeting eyewitnesses of the past in the history classroom in general. We chose this example because

she also discusses possible responses on a meta-level, rather than only her own immediate responses.

Figures 2 and 3 provide radar charts to sum up the quantities of the particular historical experiences given in the examples above. The axis displays the number of coding for each response within the example.

For the teacher, the Figure 3 indicates that the historical experience at the Hiroshima Museum in Japan is driven by the visitor's physical engagement (most codings are within this dimension) and affective engagement. The learner mentioned only one response (imagination) that referred to her cognitive engagement with the site. For the student engaging with the eyewitness, Figure 4 indicates that the experience of meeting an eyewitness in person is especially made up of the cognitive and physical dimensions. The student mentioned only one response (personal attachment) that referred to the affective engagement with the eyewitness.

The utility of these examples and the radar charts rests on their ability to show us patterns of interactions/responses within the range of elements identified that would normally be obscured by overly general descriptions of complex responses (e.g., 'the learner had a strong emotional reaction') that typifies the current literature. In this way, we are able to see how learners respond to complex historic sources and can build pedagogy or interventions that take into consideration the whole of the historical experience and the whole person experiencing it.

**Table 3**

## Hiroshima Peace Museum

Text	Coded Statements	Element	Rationale
I stood there with a <b>lump in my throat</b> . The <b>skin on my arm had turned to tiny little bumps</b> as I laid my eyes on the remnants of the preserved hair of a Hibakusha. The hair had fallen out in the weeks following the bombing and was now, immortally preserved behind a glassed case, part of a larger exhibit in the West Building of the Hiroshima Peace Memorial Museum. I was visiting the museum as part of the Japan Society's Study Tour for Educators with the intent of gathering artifacts and resources, as well as documenting the journey for my students. Yet, as I stood there staring at this straw-like bundle of hair, <b>I knew that more important than the lessons I would create or the resources I would bring back with me, was my role as a witness to this place for my students.</b>	lump in my throat	Physiological response	Visitor describes her physiological response to the exhibition piece: hair that had fallen out after the bombing of Hiroshima.
	skin on my arm had turned to tiny little bumps		
	I knew that more important than the lessons I would create or the resources I would bring back with me, was my role as a witness to this place for my students.	(Sense of) insight	The visitor understood that the experience of visiting Hiroshima Museum would somehow change her by "realizations about oneself, one's identity, and one's purpose in life" (Latham, 2013, 10).
		Awe and reverence	Visitor talks about her role as a teacher visiting the historic site and becoming a witness to history to bring back to her students. Thereby, she describes "a feeling of smallness or a sudden understanding of the grandeur in the meaning of the world around them" (Latham, 2013, 14). She refers to a sudden feeling of herself and her role as a teacher, and she acknowledges a sense of larger meaning.
Upon entering the museum, I found myself <b>stumbling through a maze-like replica of a bombed-out building</b> , passing grotesque life-like mannequins of people engulfed in flames, with their faux fire-torched skin hanging and dripping off their arms. <b>I took a moment to breathe before, with my stomach in knots and a feeling of disgust embedded deep in my gut</b> , I moved on to the first exhibit hall.	stumbling through a maze-like replica of a bombed-out building	Physical interaction	The visitor navigated herself through the site, "stumbling", indicating her interaction with the site by moving through the objects.
	took a moment to breathe... my stomach in knots and a feeling of disgust embedded deep in my gut	Physiological response	Her description of the experience clearly shows her embodied responses to the objects, referring to her bodily functions.
		Unifying experience	Her description of the situation and the resultant feeling of becoming a witness indicates a holistic experience, a moment that involves intellect, emotions and physical responses. The visitor's experiences resulted in a sense of wholeness.

Text	Coded Statements	Element	Rationale
On display here were <b>various belongings of people who were either killed or injured during the bombing</b> . Behind one glass case was the <b>personal belongings of a junior high-school student, including a cap, a belt, and a school uniform</b> . <b>“These were just kids!” I internally screamed</b> to no one in particular, as <b>my disgust slowly shifted to feelings of anger and confusion</b> .	various belongings of people who were either killed or injured during the bombing... personal belongings of a junior high-school student, including a cap, a belt, and a school uniform	Sensory interaction	The visitor is describing the objects in the exhibition: items of an innocent victim of war that are so close that they are almost touchable, “visually and spatially perceived elements... “the need to soak it up with one’s eyes”” (Latham, 2013, 10f.).
	“These were just kids!” I internally screamed... as my disgust slowly shifted to feelings of anger and confusion.	Historical proximity	Her response to the victim’s belongings of internally screaming indicates her strong feeling of temporal and spatial proximity to the place and time. For example, the teacher’s screaming may be aimed at the perpetrators, lending itself to the teacher falling back in time and place as if she was there.
		Personal attachment	Her personal attachment to the objects connects her experience at the museum through her professional identity as a high-school teacher; the proximity to a specific element of her life contributes to her strong reaction of internally screaming.
		Being moved	The visitor’s response shows her deep emotional arousal when processing the belongings of the high-school student.
		Imagination	The objects belonging to school-aged children triggered the visitor’s imagination that helped to better construct the world of the person and the circumstances that resulted in the strong response.
As I <b>slowly moved across the room</b> , I <b>could feel my muscles tense up in fear of what awaited me</b> in the next exhibit.	I slowly moved across the room	Physical interaction	The visitor describes her activity at the site, navigating her way across the room.
	I could feel my muscles tense up in fear of what awaited me	Physiological response	Her muscles tensed up, which indicates her embodied response towards what she will be confronted with next.
		Being moved	The visitor fears what she will be confronted with next. This indicates her feeling of being moved in form of the emotion fear.



Text	Coded Statements	Element	Rationale
<b>My fear was justified</b> when I was confronted with a glass-cased exhibit of a preserved part of a keloid that had been torn from someone's upper arm, alongside other bodily parts such as deformed tongues resulting from the radiation.	My fear was justified	Being moved	The exhibition objects deeply move the visitor, indicated by her description of her fear was justified.
When I reached the end of the exhibit, before I signed the visitor book intended to capture visitors' impressions of the museum, I raised <b>my water-filled eyes</b> above me to a rock engraved with the following <b>quote from Pope John Paul II upon visiting the museum in 1981: "War is the work of man. War is the destruction of man. War is death. To remember the past is to commit oneself to the future. To remember Hiroshima is to abhor nuclear war. To remember Hiroshima is to commit oneself to peace"</b> (Pope John Paul II, 1981 on display in the Hiroshima Peace Memorial Museum, 2014). <b>In that moment I felt a transformation in myself that would indubitably change the way I taught about World War II and the Bombing of Pearl Harbor. In that moment, I became an educator dedicated to morality, human rights, and social justice action.</b>	my water-filled eyes	Physiological response	The visitor describes her bodily response of tearing up at the end of the exhibition.
		Being moved	She was moved to tear by the emotional resonance of the story.
	quote from Pope John Paul II	Contextualization	The quote engraved in a rock as the closing of the exhibition triggers the visitor to understand the whole site in a bigger context, linking it to her prior knowledge.
		Awe and reverence	The quote of Pope John Paul II at the end of the experience seems to stress the meaning of the Hiroshima Museum and leaves the visitor "struck by the power" of the place. The visitor experiences deep appreciation of remembering and committing herself to peace.
	In that moment I felt a transformation in myself that would indubitably change the way I taught about World War II and the Bombing of Pearl Harbor. In that moment, I became an educator dedicated to morality, human rights, and social justice action.	(Sense of) insight	It changed her understanding of Hiroshima. It changed how she would think about (and teach) Hiroshima and Second World War. This indicates a deep understanding and even change of perspectives regarding her role as a teacher and the topic.
		Unifying experience	At the end of the exhibition, after reading the quote by Pope John Paul II, the visitor experienced the intertwining of all facets of the experience she had at the site, resulting in a sense of wholeness. The moment involved intellect, emotions and her physical responses and results in a deep understanding of her role as an educator.

Table 4

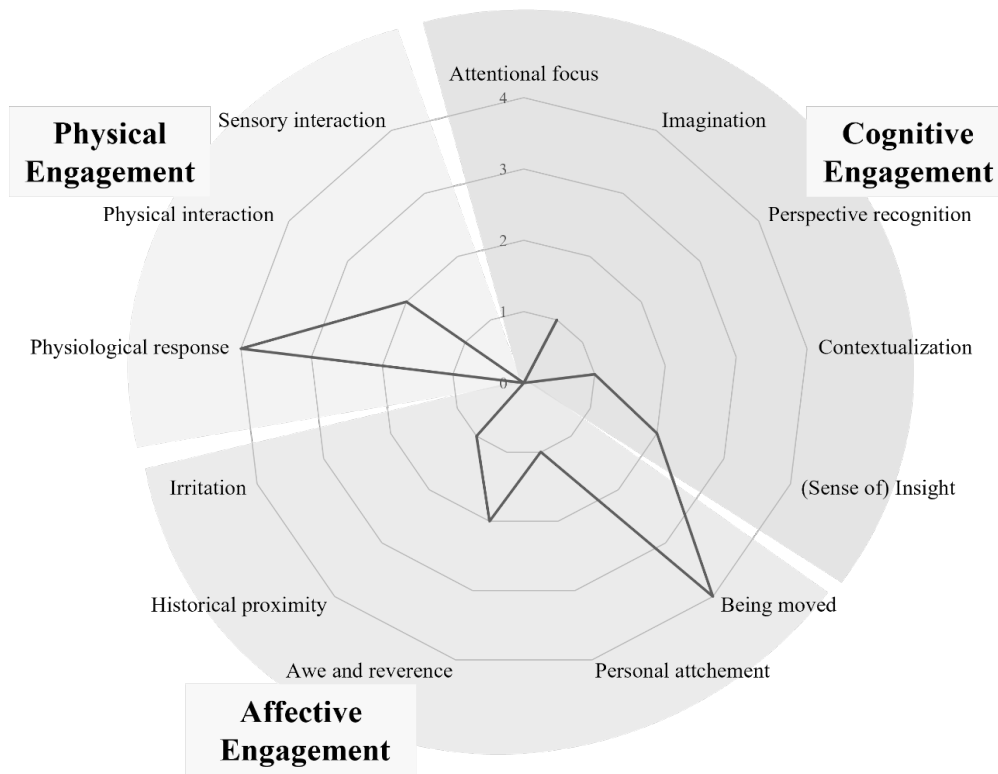
## German Democratic Republic eyewitness

Text	Coded Statements	Element	Rationale
<p><i>Interviewer:</i> How did you feel about everything? What did you enjoy? What did you find difficult?</p> <p><i>Student:</i> For example, with the posters, that was, that was a lot of time pressure and so on, but I liked it because we knew a lot about our own topic afterwards. Not about the other topics. I found that already...But that was already good. And then in between how they proceed</p> <p>I.: yes, the historians, yes.</p> <p>S.: Well, that didn't interest me so much [laughs]. I didn't really understand it either, then. And ... simply because there was so little time, I think. We couldn't do it in more detail, I think. But...<b>the eyewitness, the interview, I thought that was good, it interested me to see someone who lived there and experienced all that.</b>... But there was too little time.</p>	<p>the eyewitness, the interview, I thought that was good, it interested me to see someone who lived there and experienced all that</p>	<p>Attentional focus</p>	<p>The student explains how the fact that the witness "lived and experienced all that" triggered her interest.</p>
<p>I.: Eyewitnesses in history lessons: What do you think about an eyewitness of the past talking about his/her experiences?</p> <p>S.: Yes, as I just said, I just think: <b>you can really feel it, you can better take on the perspective if you see the person</b> and then think: he was really there, he experienced it, he knows. It's a different experience than reading a text.</p> <p>I.: You said at the beginning: You can feel it! Can you explain that a little more?</p> <p>S.: Somehow ... he then conveys, or <b>you get some kind of feeling when you hear how he talks.</b> And he then also <b>uses different voices and stresses some things, perhaps you wouldn't read that like that yourself when you read the text. And it then conveys such a certain feeling that you can't get out of the text.</b> ... That's how I see it!</p>	<p>You can really feel it, you can better take on the perspective if you see the person</p> <hr/> <p>uses different voices and stresses some things, perhaps you wouldn't read that like that yourself when you read the text. And it then conveys such a certain feeling that you can't get out of the text</p>	<p>Personal attachment</p> <hr/> <p>Sensory interaction</p>	<p>The student indicates taking on the perspective and identifying with the person and perceiving the story from his point of view.</p> <hr/> <p>By noticing the details of the eyewitness's voice while talking, the student carefully listens to the story with all its vocal nuances.</p>

<p>I.: What do you think, how should eyewitnesses appear in class: just live, as you experienced it, or as a video, or in the text?</p> <p>S.: I think live, I actually think it's the best. Because <b>you can ask the questions that interest you, and if you don't understand something, you can also ask. And you also see the reactions to the questions, which you probably don't see in a video. You also don't know what's cut out of it.</b> And so ... That's why I find it better to meet directly – that's also the atmosphere when you know he's standing in front of me and he knows exactly what we're talking about now, he's experienced it and so on.</p> <p>I.: Does it make a difference in learning?</p> <p>S.: I think so, because you remember it more when you tell it personally than when you read it as a text.</p>	<p>you can ask the questions that interest you, and if you don't understand something, you can also ask</p>	<p>Physical interaction</p>	<p>The student clearly states the advantage of being able to interact with the eyewitness by asking questions and clarifying things she did not understand. She points also to the importance of using the learning opportunity.</p>
<p>I.: Does it make a difference for you if you also know people's feelings about historical events?</p> <p>S.: Yes!</p> <p>I.: Ok, and how do you understand it?</p> <p>S.: I just think that when you read about events like this without feeling, then you interpret a lot yourself and think how you would feel there. <b>If you now hear that from someone, how he felt, then you understand the situation better, as it really was.</b> For example, if there were so many people protesting, then I would think now that they were all a bit scared and so on... or you could also think that they were in a large group, they were not scared, they wanted to be the opposition. But if you now know that they were already afraid, for example, then you see it quite differently. So you know it better, I think!</p>	<p>And you also see the reactions to the questions, which you probably don't see in a video.</p>	<p>Sensory interaction</p>	<p>The student indicates the possibility of fully observing the eyewitness while telling his/her story and the reactions to one's questions.</p>
<p>I.: Does it make a difference for you if you also know people's feelings about historical events?</p> <p>S.: Yes!</p> <p>I.: Ok, and how do you understand it?</p> <p>S.: I just think that when you read about events like this without feeling, then you interpret a lot yourself and think how you would feel there. <b>If you now hear that from someone, how he felt, then you understand the situation better, as it really was.</b> For example, if there were so many people protesting, then I would think now that they were all a bit scared and so on... or you could also think that they were in a large group, they were not scared, they wanted to be the opposition. But if you now know that they were already afraid, for example, then you see it quite differently. So you know it better, I think!</p>	<p>If you now hear that from someone, how he felt, then you understand the situation better, as it really was</p>	<p>(Sense of) insight</p>	<p>The learner states that by hearing the story of someone who was there and getting to know his/her feelings, you don't have to interpret any more but "understand the situation... as it really was".</p>
<p>I.: We always learn something for ourselves from history. If you now only have the data and facts, you also take a message with you. If you also know people's lives with their feelings, with their decisions, with their considerations and so on, you also take a message with you. What do you think is the difference for you?</p> <p>S.: Well, if I know about the people what they were thinking then, yes, I think, if I often read something like that, yes, World War II and that's how it happened, then you can't quite comprehend the decisions, yes, why they did it that way. <b>If you just know the thoughts, then... you see that much more logically and understand it also why it was like that and then you can remember it better.</b> If you think that it happened because he thought that and that...</p>	<p>If you just know the thoughts, then... you see that much more logically and understand it also why it was like that</p>	<p>Contextualization</p>	<p>The student explains that getting to know the thoughts of the historical agent helps to understand the past events more fully and situate them in a broader context as compared to just reading a text about the event.</p>
<p>I.: We always learn something for ourselves from history. If you now only have the data and facts, you also take a message with you. If you also know people's lives with their feelings, with their decisions, with their considerations and so on, you also take a message with you. What do you think is the difference for you?</p> <p>S.: Well, if I know about the people what they were thinking then, yes, I think, if I often read something like that, yes, World War II and that's how it happened, then you can't quite comprehend the decisions, yes, why they did it that way. <b>If you just know the thoughts, then... you see that much more logically and understand it also why it was like that and then you can remember it better.</b> If you think that it happened because he thought that and that...</p>	<p>and then you can remember it better</p>	<p>Attentional focus</p>	<p>The student offers a meta-analysis that you remember the information given within an eyewitness's story better than reading a factual text about an event, indicating the opportunity that the personal interaction creates deeper concentration on the learning material.</p>

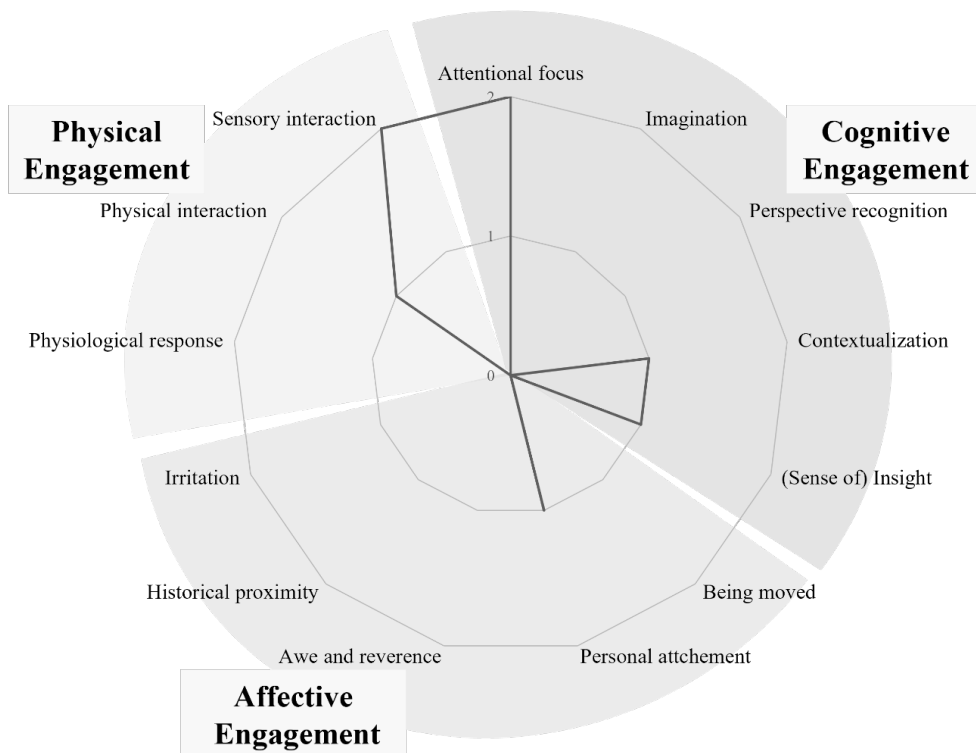
**Figure 2**

Hiroshima Museum showing the coding frequency for each response



**Figure 3**

Coding frequency for each response for the student interview about the eyewitness encounter



---

### **Conclusion, limitations and outlook**

We developed a conceptual model to describe and define the learning experience with complex historical sources—namely encountering eyewitness of the past and visiting historic sites. We drew upon empirical, theoretical, and conceptual works from a range of related disciplines including media studies, museology, anthropology, cultural geography, and educational psychology, to develop a deeper understanding of what is part of this complex experience and how we can grasp it. We argue that embedded within perceiving these sources as authentic, learners cognitively, affectively and physically engage with the learning material. In the same way that we identify a range of possible interconnected responses that can occur as part of the historical experience, the unifying experience is possible, but it is neither required, nor guaranteed.

The framework is the first endeavor to conceptually encapsulate the learning experience with complex historical sources. In particular, our model contributes to a better understanding of the complex and intertwining responses involved in learning about the past. However, we recognize the disequilibrium in a model that delineates far more cognitive elements and fewer physical ones. This is a result of the extant state of the literature, rather than a determination that there are fewer physical considerations for learners. We offer this model as the beginning of the conversation about how to understand learning with complex historical sources, not the ending. We invite fellow researchers to test this model, qualitatively and quantitatively, to deepen our understanding of the elements we have identified and further delineate underlying processes and mechanisms of these encounters. At this point, we do not know if response to one element triggers or is a prerequisite for another. Nor do we know if certain elements cluster together, and, if so, under what conditions. Further, we recognize that there are differences between encounters with historic sites and encounters with eyewitnesses that may become sharper as researchers engage with this framework. The more this model is used and refined, the more it may reveal further complexities in the intertwining dynamics of historical experiences that contribute to a better understanding of the learning processes behind these experiences.

---

### References

- Angvik, M. and von Borries, B. (Eds.) (1997). *Youth and history. A comparative European survey on historical consciousness and political attitudes among adolescents*. Hamburg: Körber-Stiftung.
- Arnold-de Simine, S. (2012). Memory museum and museum text: Intermediality in Daniel Libeskind's Jewish Museum and W. G. Sebald's Austerlitz. *Theory and Culture*, 29(1), 14–35.
- Baron, C. (2012). Understanding historical thinking at historic sites. *Journal of Educational Psychology*, 104(3), 833–847. <https://doi.org/10.1037/a0027476>
- Baron, C. (2013). Using inquiry based instruction to encourage teachers' historical thinking at historic sites. *Teaching and Teacher Education*, 35, 157–169.
- Baron, C. (2016). Using embedded visual coding to support contextualization of historical texts. *American Educational Research Journal*, 53(3), 516–540. <https://doi.org/10.3102/0002831216637347>
- Baron, C., Sklarwitz, S., Bang, H. and Shatar, H. (2019). Understanding what teachers gain from professional development at historic sites. *Theory & Research in Social Education*, 47(1), 76–107.
- Barton, K. C. and Levstik, L. (2004). *Teaching history for the common good*. Mahwah, NJ: Laurence Erlbaum.
- Benjamin, W. (1965). *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. Frankfurt am Main: Suhrkamp.
- Bertram, C., Wagner, W. and Trautwein, U. (2017). Learning historical thinking with oral history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54(3), 444–484. <https://doi.org/10.3102/0002831217694833>
- Bilandzic, H. and Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, 36(1), 29–50. <https://doi.org/10.1515/COMM.2011.002>
- Boler, M. (1999). *Feeling power: Emotions and education*. New York: Routledge.
- Brauer, J. (2016a). Empathy as an emotional practice in historical pedagogy. *Miscellanea Anthropologica et Sociologica*, 17(4), 27–44.
- Brauer, J. (2016b). ‘Heiße Geschichte’? Emotionen und historisches Lernen in Museen und Gedenkstätten. In S. Willner, G. Koch and S. Samida (Eds.), *Doing History - Performative Praktiken in der Geschichtskultur* (pp. 29–44). Münster: Waxmann.

- 
- Brauer, J. and Lücke, M. (Eds.) (2013). *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven*. Göttingen: V&R unipress.
- Busselle, R. and Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, 18(2), 255–280. <https://doi.org/10.1111/j.1468-2885.2008.00322.x>
- Cameron, C. M. and Gatewood, J. B. (2003). Seeking numinous experiences in the unremembered past. *Ethnology*, 42(1), 55–71.
- Cobb, P. and Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 28(2), 4–15.
- Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, 4, 245–264. [https://doi.org/10.1207/S15327825MCS0403\\_01](https://doi.org/10.1207/S15327825MCS0403_01)
- Csikszentmihalyi, M. and Robinson, R. E. (1990). *The art of seeing: An interpretation of the aesthetic encounter*. Los Angeles: Getty Center for Education in the Arts.
- Dawson, E. (2014). Equity in informal science education: developing an access and equity framework for science museums and science centers. *Studies in Science Education*, 50(2), 209–247.
- De Leur, T., Van Boxtel, C. and Wilschut, A. (2017). ‘I saw angry people and broken statues’: Historical empathy in secondary history education. *British Journal of Educational Studies*, 65(3), 331–352. <https://doi.org/10.1080/00071005.2017.1291902>
- Dierking, L. D. and Falk, J. H. (1992). *The museum experience*. Washington, D.C.: Howells House.
- Drozdowski, D., De Nardi, S. and Waterton, E. (2016). Geographies of memory, place, and identity: Intersections in remembering war and conflict. *Geography Compass*, 10(11), 447–456. <https://doi.org/10.1111/gec3.12296>
- Dutt-Doner, K. M., Allen, S. and Campanaro, K. (2016). Understanding the impact of using oral histories in the classroom. *The Social Studies*, 107(6), 257–265. <https://doi.org/10.1080/00377996.2016.1221792>
- Ellsworth, E. (2005). *Places of learning: Media, architecture, pedagogy*. New York: Routledge.
- Endacott, J. L. (2010). Reconsidering affective engagement in historical empathy. *Theory and Research in Social Education*, 38(1), 6–47. <https://doi.org/10.1080/00933104.2010.10473415>

- Endacott, J. L. and Brooks, S. (2013). An updated theoretical and practical model for promoting historical empathy. *Social Studies Research and Practice*, 8(1), 41–58.
- Endacott, J. L. and Brooks, S. (2018). Historical empathy: Perspectives and responding to the past. In S. Metzger and L. McArthur Harris (Eds.), *The Wiley International Handbook of History Teaching and Learning*. New York: Wiley-Blackwell.
- Fong, T., Nourbakhsh, I. and Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3–4), 143–166.
- Frank, B. (2015). *Presence messen in laborbasierter Forschung mit Mikrowelten*. Wiesbaden: Springer.
- Gatewood, J. B. and Cameron, C. M. (2004). Battlefield pilgrims at Gettysburg National Military Park. *Ethnology*, 43(3), 193–216.
- Glaser, M. and Schwan, S. (2015). Explaining pictures: How verbal cues influence processing of pictorial learning material. *Journal of Educational Psychology*, 107(4), 1006–1018. <https://doi.org/10.1037/edu000004>
- Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, 38(2), 247–266. <https://doi.org/10.1207/s15326950dp3802>
- Green, M. C. and Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Greene, J., Hitt, C., Kraybill, A. and Bogulski, C. A. (2015). Learning from live theatre. *Education Next*, 54–61.
- Greeno, J. G. and Engstrom, Y. (2014). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 128–150). Cambridge University Press.
- Grever, M. (2018). Teaching the war: Reflections on popular uses of difficult heritage. In T. Epstein and C. L. Peck (Eds.), *Teaching and learning difficult histories in international contexts; A critical sociocultural approach* (pp. 30–44). New York: Routledge.
- Gussmann, M., Merkt, M. and Schwan, S. (2017). *Impact of authentic locations*. Paper presented at the Conference *The Concept of Aura*, University of Konstanz, Konstanz, Baden-Württemberg, Germany. December 14-18, 2017.
- Hampp, C. and Schwan, S. (2014). The role of authentic objects in museums of the history of science and technology: Findings from a visitor study. *International Journal of*



- 
- Science Education, Part B*, 5(2), 161–181.  
<https://doi.org/10.1080/21548455.2013.875238>
- Hartmann, U. and Hasselhorn, M. (2008). Historical perspective taking: A standardized measure for an aspect of students' historical thinking. *Learning and Individual Differences*, 18(2), 264–270. <https://doi.org/10.1016/j.lindif.2007.10.002>
- Hooper-Greenhill, E. (2004). Measuring learning outcomes in museums, archives and libraries: The learning impact research project (LIRP). *International Journal of Heritage Studies*, 10(2), 151–174. <https://doi.org/10.1080/13527250410001692877>
- Huijgen, T., Van Boxtel, C., van de Grift, W. and Holthuis, P. (2017). Toward historical perspective taking: Students' reasoning when contextualizing the actions of people in the past. *Theory and Research in Social Education*, 45(1), 110–144.  
<https://doi.org/10.1080/00933104.2016.1208597>
- Hutchins, E. (1995). *Cognition in the wild* (No. 1995). Cambridge, MA: MIT Press.
- Jones, S. (2016). Unlocking essences and exploring networks: Experiencing authenticity in heritage education settings. In C. Van Boxtel, M. Grever and S. Klein (Eds.), *Sensitive pasts: Questioning heritage in education* (pp. 130–152). New York: Berghahn.
- Joy, A. and Sherry, J. F. (2003). Speaking of art as embodied imagination: A multisensory approach to understanding aesthetic experience. *Journal of Consumer Research*, 30, 259–282.
- Keltner, D. and Haidt, J. (2003). Approaching awe, a moral, spiritual and aesthetic emotion. *Cognition & Emotion*, 17(2), 297–314. <https://doi.org/10.1080/02699930244000318>
- Kirk, D. and Kinchin, G. (2003). Situated learning as a theoretical framework for sport education. *European Physical Education Review*, 9(3), 221–235.  
<https://doi.org/10.1177/1356336X030093002>
- Klimmt, C. and Vorderer, P. (2003). Media psychology “is not yet there”: Introducing theories on media entertainment to the presence debate. *Presence: Teleoperators and Virtual Environments*, 12(4), 346–359. <https://doi.org/10.1162/105474603322391596>
- Kohlmeier, J. (2006). “Couldn't she just leave?": The relationship between consistently using class discussions and the development of historical empathy in a 9<sup>th</sup> grade world history course. *Theory and Research in Social Education*, 24(1), 34–57.
- Korthagen, F. A. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*, 26(1), 98–106. <https://doi.org/10.1016/j.tate.2009.05.001>

- Krämer, S. (2012). Zum Paradoxon von Zeugenschaft im Spannungsfeld von Personalität und Depersonalisierung. In M. Rössner and H. Uhl (Eds.), *Renaissance der Authentizität? Über die neue Sehnsucht nach dem Ursprünglichen* (pp. 15–26). Bielefeld: transcript.
- Kunter, M. and Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: UTB.
- Landström, S., Granhag, P. A. and Hartwig, M. (2005). Witnesses appearing live versus on video: Effects on observers' perception, veracity assessments and memory. *Applied Cognitive Psychology*, 19(7), 913–933. <https://doi.org/10.1002/acp.1131>
- Latham, K. F. (2007). The poetry of the museum: A holistic model of numinous museum experiences. *Museum Management and Curatorship*, 22(3), 247–263. <https://doi.org/10.1080/09647770701628594>
- Latham, K. F. (2013). Numinous experiences with museum objects. *Visitor Studies*, 16(1), 3–20. <https://doi.org/10.1080/10645578.2013.767728>
- Lee, J., Hicks, D., Henriksen, D., Mishra, P. and Cain, W. (2015). Historical soundscapes for creative synthesis. *TechTrends*, 59(5), 4–8.
- Lee, P. J. (1984). Historical imagination. In A. K. Dickinson, Lee, P.J., and Rogers, P.J. (Eds.), *Learning History* (pp. 85–116). London: Heinemann Educational Books.
- Lee, P. and Ashby, R. (2001). Empathy, perspective taking, and rational understanding. In O. L. Davis, E. A. Yeager and S. J. Foster (Eds.), *Historical empathy and perspective taking in the social studies* (pp. 21–50). Lanham: Rowman & Littlefield.
- Lombard, M. and Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2), 1–36.
- Manzano, Ö., Theorell, T., Harmat, L. and Ullén, F. (2010). The psychophysiology of flow during piano playing. *Emotion*, 10(3), 301–311. <https://doi.org/10.1037/a0018432>
- Maiese, M. (2014). How can emotions be both cognitive and bodily? *Phenomenology and the Cognitive Sciences*, 13(4), 513–531.
- Mason, R., Galani, A., Lloyd, K. and Sayner, J. (2018). Experiencing mixed emotions in the museum: empathy, affect, and memory in visitor's responses to histories of migration. In L. Smith, M. Wetherell, & G. Campbell, (Eds), *Emotion, affective practices, and the past in the present* (pp. 124–148). London: Routledge.
- Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M. and Jacobsen, T. (2015). Towards a psychological construct of being moved. *PLoS ONE*, 10(6), 33–35. <https://doi.org/10.1371/journal.pone.0128451>

- 
- Muth, C. and Carbon, C. C. (2013). The aesthetic aha: On the pleasure of having insights into Gestalt. *Acta Psychologica*, 144(1), 25–30.  
<https://doi.org/10.1016/j.actpsy.2013.05.001>
- Nacke, L. E. and Lindley, C. A. (2008). Flow and immersion in first-person shooters: Measuring the player's gameplay experience. *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, 81–88. <https://doi.org/10.1145/1496984.1496998>
- Nardi, B. A. (Ed.). (1996). *Context and consciousness: Activity theory and human-computer interaction*. Cambridge, MA: MIT Press.
- Perry, D. L. (2002). Profound learning: Stories from museums. *Educational Technology*, 42(2), 21–25.
- Plato, A. von (2009). Contemporary witnesses and the historical profession: Remembrance, communicative transmission, and collective memory in qualitative history. *Oral History Forum d'histoire orale*, 29, 1–27. Retrieved from <http://www.oralhistoryforum.ca/index.php/ohf/article/view/56>
- Ramlogan, S., Raman, V. and Sweet, J. (2014). A comparison of two forms of teaching instruction: Video vs. live lecture for education in clinical periodontology. *European Journal of Dental Education*, 18(1), 31–38. <https://doi.org/10.1111/eje.12053>
- Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, 30(1), 86–112.
- Reiß, M., Becker, A. and Krist, H. (2014). Gibt es einen Videodefiziteffekt bei Aufgaben zur Theory of Mind? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 46(3), 155–163. <https://doi.org/10.1026/0049-8637/a000112>
- Resnick, L. B. (1987). *Education and learning to think*. Washington, D.C.: National Academies.
- Rose, J. (2016). *Interpreting difficult history at museums and historic sites*. New York: Rowman & Littlefield.
- Rosenberg, E. (2007). The geography of memory: Walking as remembrance. *The Hedgehog Review*, (Summer), 54–67.
- Rosenzweig, R. and Thelen, D. P. (1998). *The presence of the past: Popular uses of history in American life*. New York: Columbia University Press.
- Rössner, M., and Uhl, H. (2012). *Renaissance der Authentizität. Über die neue Sehnsucht nach dem Ursprünglichen*. Bielefeld: transcript.

- Sabrow, M. (2012). Der Zeitzeuge als Wanderer zwischen den Welten. In M. Sabrow & N. Frei (Eds.), *Die Geburt des Zeitzegen nach 1945* (pp. 13–32). Göttingen: Wallstein Verlag.
- Sakr, M., Jewitt, C. and Price, S. (2016). Mobile experiences of historical place: A multimodal analysis of emotional engagement. *Journal of the Learning Sciences*, 25(1), 51–92. <https://doi.org/10.1080/10508406.2015.1115761>
- Savenije, G. M. (2016). An intriguing historical trace or heritage? Learning about another person's heritage in an exhibition addressing the Second World War. In C. Van Boxtel, M. Grever and S. Klein (Eds.), *Sensitive pasts: Questioning heritage in education* (pp. 218–239). New York: Berghahn Books.
- Savenije, G. M. and De Bruijn, P. (2017). Historical empathy in a museum: Uniting contextualization and emotional engagement. *International Journal of Heritage Studies*, 23(9), 1–14.
- Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M. and Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLoS ONE*, 12(6), 1–45. <https://doi.org/10.17605/OSF.IO/Q8ZV5>
- Schörken, R. (1998). Imagination und geschichtliches Verstehen, Neue Sammlung. *Vierteljahres-Zeitschrift für Erziehung und Gesellschaft*, 38(2), 202–212.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Thrash, T. M. and Elliot, A. J. (2004). Inspiration: Core characteristics, component processes, antecedents, and function. *Journal of Personality and Social Psychology*, 87(6), 957–973. <https://doi.org/10.1037/0022-3514.87.6.957>
- Toila-Kelly, D. P. (2018). Race and affect at the museum. In D. P. Toila-Kelly, E. Waterton, and S. Watson (Eds.), *Heritage, affect, and emotion; Politics, practices, and infrastructures* (pp.33–46). New York: Routledge.
- Trofanenko, B. (2006). Interrupting the gaze: On reconsidering authority in the museum. *Journal of Curriculum Studies*, 38, 49–65. [10.1080/00220270500038511](https://doi.org/10.1080/00220270500038511).
- Trope, Y. and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463. <https://doi.org/10.1037/a0018963>
- Uhl, H. (2012). Orte und Lebenszeugnisse - “Authentizität” als Schlüsselkonzept in der Vermittlung der NS-Verfolgungs- und Vernichtungspolitik. In M. Rössner and H. Uhl

- (Eds.), *Renaissance der Authentizität? Über die neue Sehnsucht nach dem Ursprünglichen* (pp. 258–284). Bielefeld: transcript.
- Van Boven, L., Kane, J., McGraw, A. P. and Dale, J. (2010). Feeling close: Emotional intensity reduces perceived psychological distance. *Journal of Personality and Social Psychology*, 98(6), 872–885. <https://doi.org/10.1037/a0019262>
- Van Boxtel, C, Grever, M. and Klein, S. (2016). Introduction: The appeal of heritage education. In C. Van Boxtel, M. Grever and S. Klein (Eds.), *Sensitive pasts; Questioning heritage in education* (pp. 1–18). New York: Berghahn.
- Van Nieuwenhuysse, K. (2011). My history begins in Africa. In C. Van Boxtel, S. Klein and E. Snoep (Eds.), *Heritage education: Challenges in dealing with the past* (pp. 26–28). Amsterdam: Erfgoed Nederland.
- Van Drie, J. and Van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, 20(2), 87–110. <https://doi.org/10.1007/s10648-007-9056-1>
- Vogl, E., Pekrun, R., Murayama, K. and Loderer, K. (2019). Surprised - curious - confused: Epistemic emotions and knowledge exploration. *Emotion*, 20(1), 625–241. <https://doi.org/10.1037/emo0000578>
- Wetherell, L., Smith, M. and Cambell, G. (2018). Introduction: Affective heritage practices. In L. Wetherell, M. Smith and G. Cambell (Eds.), *Emotion, affective practices and the past in the present* (pp. 1–21). London: Routledge.
- Waterton, E. (2018). Curating affect: Exploring the historical geography-heritage studies nexus at Sovereign Hill. *Australian Geographer*, 49(1), 219–235.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636. doi:10.3758/BF03196322
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>
- Wineburg, S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science*, 22(3), 319–346.
- Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Philadelphia, PA: Temple University.
- Zembylas, M. (2007). Theory and methodology in researching emotions in education. *International Journal of Research & Method in Education*, 30(1), 57–72. doi:10.1080/17437270701207785

- Zembylas, M. (2016). Making sense of the complex entanglement between emotion and pedagogy: Contributions of the affective turn. *Cultural Studies of Science Education*, 11(3), 539–550.
- Zembylas, M. (2018). Understanding the emotional regimes of reconciliation in engagements with ‘difficult heritage’. In L. Smith, M. Wetherell and G. Campbell (Eds.), *Emotion, affective practices and the past in the present* (pp. 197–212). London: Routledge

---

# 4

## STUDIE 2

---

Zachrich, L., Wagner, W., Bertram, B. & Trautwein, U. (2022). Das „Aura-Erlebnis“ in der Begegnung mit Zeitzeug\*innen – Eine komplexe Lernerfahrung verstehen und standardisiert erfassen In M. Nitsche & M. Waldis (Hrsg.), *Geschichtsdidaktisch intervenieren*. Bern: hep-Verlag.  
*Manuskript akzeptiert zur Publikation.*

**Abstract**

Der Begegnung mit Zeitzeug\*innen wird von Schüler\*innen eine besondere Bedeutung zugeschrieben. Der vorliegende Beitrag beschreibt die Entwicklung eines standardisierten Fragebogens zur Erfassung der pädagogisch-psychologischen Prozesse des „Aura-Erlebnisses“ in der Begegnung mit Zeitzeug\*innen sowie deren Zusammenhänge zu Situations- und Personenmerkmalen. Die Datengrundlage für diese Validierungsstudie bildeten 668 Schüler\*innen der Klassen 8 bis 13 aus Baden-Württemberg und Brandenburg, die Zeitzeug\*innen aus der ehemaligen DDR begegnet sind. Wie theoretisch angenommen, konnten wir die Lernerfahrung über Prozesse der kognitiven und emotionalen Involviertheit empirisch erfassen. Allerdings zeigten sich auch Abweichungen zum theoretischen Modell: Die physische Involviertheit konnten wir empirisch nicht abbilden. Wir fanden Hinweise auf die Bedeutsamkeit der Personenmerkmale, nicht jedoch der Situationsmerkmale. Der standardisierte Fragebogen ist nun für weitere Studien einsatzfähig.



### Einleitung

„Man spürt es vielmehr [...], man kann sich besser reinversetzen, wenn man den [Zeitzeugen] so sieht, und dann denkt, der war da wirklich, er hat es mitbekommen, er weiß das. Es ist ein anderes Erlebnis, als wenn man ‘nen Text liest.“<sup>1</sup> So beschreibt eine Schülerin der 9. Klasse ihre Begegnung mit einem Zeitzeugen aus der ehemaligen DDR im Geschichtsunterricht. Tatsächlich stellt die Begegnung mit Zeitzeug\*innen, so die Beobachtung von Obens und Geißler (2010), für Schüler\*innen oft ein sehr bedeutsames Erlebnis dar, das sie kaum in Worte fassen können. Aber auch für die Geschichtsdidaktik ist gar nicht so einfach zu fassen, was die Begegnung mit Zeitzeug\*innen so besonders macht: „In der Fachdebatte bedienen sich Experten deshalb häufig ‚Hilfskonstruktionen‘ wie ‚Aura‘ oder ‚Authentizität‘, ohne diese im Detail zu beschreiben“ (Obens & Geißler, 2010, 59). Ein zentrales Charakteristikum besteht möglicherweise darin, dass Zeitzeug\*innen das Potenzial haben, „nicht nur Köpfe, sondern auch Herzen der Jugendlichen zu erreichen“ (Uhl, 2012, 59). Im Sinne Pestalozzis hieße das, dass die Begegnung mit Zeitzeug\*innen neben Gedanken auch die Gefühle der Jugendlichen ansprechen kann (Klika, 2018, 79). Dass das Lernen über die Geschichte „etwas mit Emotionen zu tun“ (Pandel, 2012, 148) hat, wird zunehmend in geschichtsdidaktischen Diskussionen thematisiert (z. B. Brauer & Lücke, 2013). Bislang ist jedoch umstritten, ob Emotionen eher „Katalysator“ oder „Sackgasse“ für das Lernen der Schüler\*innen im Geschichtsunterricht sind (Brauer, 2019, 278).

Die Lernerfahrung mit Zeitzeugenberichten stellt eine Unterrichtssituation dar, die diesen Zwiespalt – Katalysator oder Sackgasse – in besonderer Weise verdeutlicht. Das Ziel des Historisch-Denken-Lernens beinhaltet unter anderem, Quellen kritisch hinterfragen zu können, um ein (selbst-)reflexives Geschichtsbewusstsein zu entwickeln (z. B. Schreiber, 2002, 27). Zeitzeugenberichte haben die Chance, Schüler\*innen für Geschichte zu motivieren, indem sie die Vergangenheit „lebendig und authentisch erfahrbar machen“ (Henke-Bockschatz, 2007, 354). Mit Blick auf das Ziel des Geschichtsunterrichts bergen sie allerdings auch das Risiko, zu überwältigen. Denn die Arbeit mit Zeitzeug\*innen stelle nach Wierling (1997) „hohe Ansprüche an das Einfühlungsvermögen und die Quellendistanz“ (239). Diese Ambivalenz der Zeitzeugenbegegnungen konnten Bertram, Wagner und Trautwein (2017, 472–473) empirisch bestätigen. Schüler\*innen berichteten nach einer Live-Begegnung mit einem\*r Zeitzeug\*in, dass sie mehr Spaß im Unterricht hatten, zeigten aber weniger Einsicht in den

---

<sup>1</sup> Das Zitat stammt aus einem Interview mit einer Schülerin nach der Unterrichtsstunde mit einem Zeitzeugen im Rahmen einer Interventionsstudie zur Wirksamkeit von Zeitzeugenbefragungen (Bertram, Wolfgang & Trautwein, 2017).

Konstruktcharakter von Geschichte als Schüler\*innen, die mit videografierten oder schriftlichen Zeugnissen lernten.

Was macht also die Live-Begegnung mit einem\*r Zeitzeug\*in besonders? Bertram und Kollegen boten zur möglichen Erklärung ihrer Befunde das von Sabrow (2012) geprägte Konzept der „Aura der Authentizität“ (27; Bertram et al., 2017, 472) an. Zachrich, Weller, Baron und Bertram (2020, 249) entwickelten zur Annäherung an diese Besonderheit ein konzeptuelles Modell, das die Merkmale der Lernerfahrung mit komplexen historischen Quellen (u.a. Zeitzeugenberichte) beschreibt und im Kern ein Zusammenspiel von kognitiver, emotionaler und physischer Involviertheit identifiziert. Die empirische Frage, was Kognition, Emotion und Physis in diesem Kontext ausmacht, ist bisher jedoch unbeantwortet. Dieser Frage stellen wir uns im vorliegenden Beitrag: Wir haben einen standardisierten Fragebogen zur Besonderheit der Begegnung mit einem\*r Zeitzeug\*in entwickelt und berichten hier Befunde zu dessen psychometrischer Qualität sowie zu Zusammenhängen mit Situations- und Personenmerkmalen.

### **Theoretischer Hintergrund**

#### **Zeitzeugenbegegnungen im Kontext historischen Lernens**

Das Verständnis von Zeitzeug\*innen unterscheidet sich in Geschichtswissenschaft, -didaktik und Erinnerungskultur (Bertram, 2020, 259–260). Dabei hilft es, die Zeitzeugenschaft in ihrer Komplexität wahrzunehmen und zu verstehen. Je nach Zielsetzung der Bildungsarbeit werden unterschiedliche Zeitzeugnisse genutzt: In der Aufarbeitung von und Erinnerung an Gewalt- und Unrechtsgeschichte sollen die Zeugnisse von Held\*innen oder Opfern für die Geschichte sensibilisieren (von Plato, 2001, 136). In der sogenannten Oral History hingegen, die sich als Methode der Geschichtswissenschaft seit den 1950er/1960er Jahren etabliert hat, geht es vor allem darum, jenen Bevölkerungsgruppen eine Stimme zu geben, die in der traditionellen Geschichtsschreibung nicht gehört wurden und werden. Es sind also Alltagsgeschichten, die Einblicke in die vergangene Zeit geben sollen (z. B. Lanman & Wendling, 2006). Zeitzeug\*innen unterscheiden sich folglich hinsichtlich verschiedener Dimensionen, z. B. je nach Rolle in der Vergangenheit (z. B. Täter, Opfer, ...), nach ihrer Expertise (Laie vs. Experte) und in dem zeitlichen Abstand, den sie zu ihren Erfahrungen in der Vergangenheit haben. Entsprechend unterscheiden sich die Berichte und auch die Begegnungen stark. Was den Berichten aber gemein ist: Sie ermöglichen Einblicke in die individuelle Wahrnehmung vergangener Erfahrungen einer Person. Diese natürlich subjektiven, aber in ihrer Erlebnisqualität „echten“ Geschichten haben das Potenzial, Schüler\*innen anders

anzusprechen als Schulbuchtexte und das historische Denken (u.a. FUER Kompetenz-Strukturmodell, Schreiber et al., 2006) zu fördern (Henke-Bockschatz, 2007, 363).

Empirisch gibt es bisher allerdings nur wenige Erkenntnisse darüber, was und wie Schüler\*innen mit und von den Begegnungen lernen. Dass Jugendliche den Zeitzeugnissen mehr Vertrauen schenken als den Historiker\*innen, war ein Ergebnis der Studie *Youth & History* (Angvik & von Borries, 1997). Obens und Geißler (2010, 59) zogen aus ihren qualitativen Interviews mit Schüler\*innen nach einer Begegnung mit Holocaust-Überlebenden das Fazit, dass viele der Jugendlichen nicht viel mehr aus den Gesprächen mitnahmen als eine emotionale Ergriffenheit. Auch Bertram et al. (2017) bestätigten mit ihrer Studie die Risiken und Potenziale der Zeitzeugenbegegnungen: In einer groß angelegten Interventionsstudie mit 962 Schüler\*innen aus 35 Klassen an weiterführenden Schulen in Baden-Württemberg untersuchten die Autor\*innen die Effekte einer Live-Begegnung mit einem\*r Zeitzeug\*in der ehemaligen DDR im Vergleich zu Video- und Textversionen der Zeitzeugenberichte auf historische Sachkompetenzen und die Motivation der Schüler\*innen. Lernende, die dem\*der Zeitzeug\*in live begegnet waren, berichteten zwar mehr Spaß an der Unterrichtseinheit, zeigten aber eine weniger ausgeprägte Einsicht in die Perspektivität von Zeitzeugenaussagen als Lernende, die mit Text- bzw. Videoberichten arbeiteten (Bertram et al., 2017, 471–473). Brüning (2018, 372) schlussfolgerte in ihrer Studie, dass die Zeugnisse von Holocaust-Überlebenden, selbst wenn die Schüler\*innen nur mit den Videos arbeiteten, sich für eine methodische Quellenkritik nicht eigneten. Es brauche eine angemessene Methodik im Umgang mit diesen besonderen Zeitzeugnissen.

### **Prozesse der Lernerfahrung**

Eine notwendige Voraussetzung für die Entwicklung einer angemessenen Methodik ist, dass wir die besondere Lernerfahrung zunächst genauer verstehen. Diesem Forschungsdesiderat haben sich Zachrich et al. (2020) gewidmet und aus mehreren theoretischen Ansätzen ein Modell entwickelt, das beschreibt, was die Erfahrung mit komplexen historischen Quellen<sup>2</sup> ausmacht – was also hinter der „Aura der Authentizität“ (Bertram, 2017, 473; Sabrow, 2012, 27) steckt (Zachrich et al., 2020, 244). Komplexe historische Quellen zeichnet aus, dass sie sich aus mehreren Medien und/oder Interaktionsebenen zusammensetzen und die Rezipient\*innen auf kognitive, emotionale und physische Weise ansprechen (Zachrich et al., 2020). Zur Beschreibung der besonderen

---

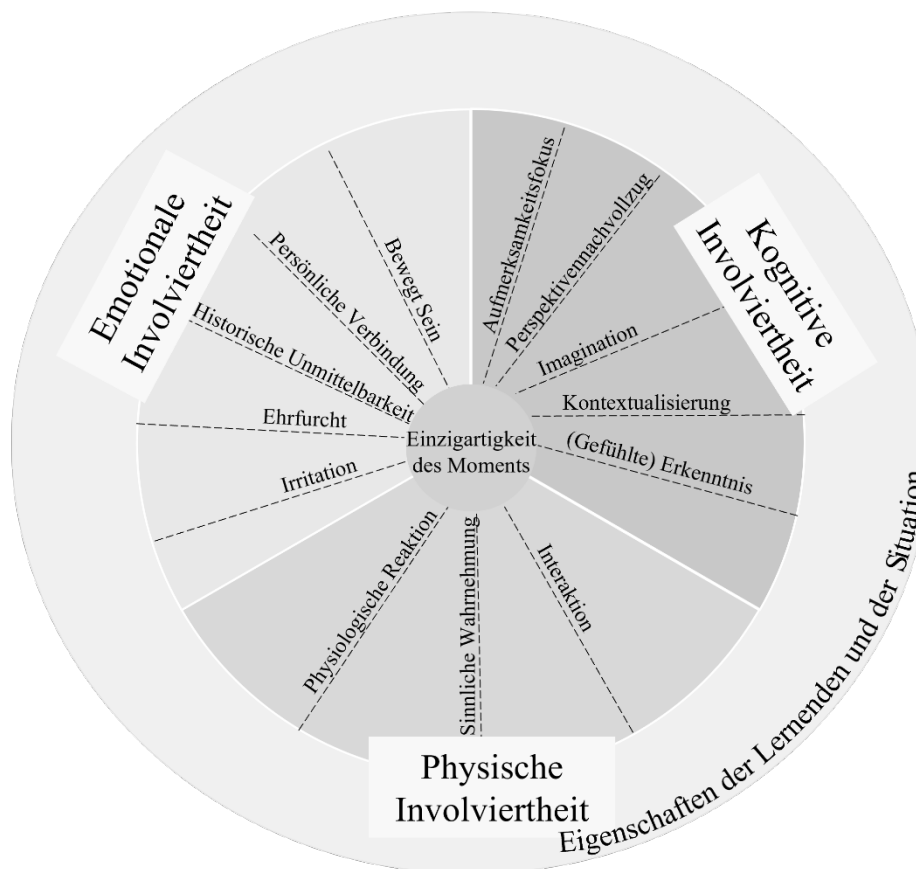
<sup>2</sup> Zeitzeugenberichte sind Quelle und Darstellung zugleich. Als Artefakt der Vergangenheit lässt sich der Bericht als Quelle dekonstruieren. Die Erzählung beinhaltet aber immer auch die eigene Interpretation der Erfahrung, dynamische Erinnerungen, und ist folglich auch Darstellung (z. B. Bertram, 2020).

Erfahrung bedienten sich die Autor\*innen mehrerer Theorien aus verschiedenen wissenschaftlichen Disziplinen: Die *Transportation Theory* (Green & Brock, 2000) und das *Model of Narrative Comprehension and Engagement* (Busselle & Bilandzic, 2009) kommen aus der Medienpsychologie und beschreiben das kognitive und emotionale Eintauchen in einen Medieninhalt, bei dem alle mentalen Kapazitäten auf die Inhalte des Narrativs gerichtet sind; die ästhetische Begegnung mit Kunstwerken beschreibt aus anthropologischer Perspektive ein Rahmenmodell zu den Motiven und Prozessen der Rezeption von Kunst (Csikszentmihalyi & Robison, 1990); Cameron und Gatewood (2003) sowie Latham (2013) identifizierten, ebenfalls aus anthropologischer Perspektive, mit dem Konzept *Numen-seeking* ein Motiv für den Besuch von Gedenkstätten, das eine nahezu spirituelle Begegnung beinhaltet; *Historische Empathie* ist ein Konzept aus der US-amerikanischen Geschichtsunterrichtsforschung und beschreibt die kognitiven und affektiven Komponenten von Empathie mit einem historischen Akteur; *Presence*, ein Konzept aus Psychologie und Informatik meint ein Gefühl, in einer medienvermittelten Welt zu sein, ohne das Medium wahrzunehmen (Lombard & Ditton, 1997). In Anlehnung an die Ausführungen dieser Konzepte entwickelten die Autor\*innen das Modell und generierten die einzelnen Facetten. Dabei verknüpften sie die theoretischen Darstellungen mit empirischen Ergebnissen aus ihren Forschungsarbeiten.

Abbildung 1 zeigt das konzeptuelle Modell der Begegnung mit komplexen historischen Quellen. Im Kern der Begegnung stehen die kognitive, emotionale und physische Involviertheit der Lernenden, die durch die Begegnung hervorgerufen werden (bzw. werden kann; Zachrich et al., 2020, 252–253). Das Modell bezieht sich beispielhaft auf Zeitzeugenbegegnungen und historische Orte. Die Lernerfahrungen können sich natürlich je nach Zeugnis und Quelle unterscheiden. Insbesondere in der physischen Involviertheit sind Unterschiede zwischen den Begegnungen denkbar. So sitzen die Schüler\*innen im Gespräch mit Zeitzeug\*innen in der Regel an einem festen Platz, während sie sich am historischen Ort bewegen. Das Modell zeigt jedoch die möglichen Verarbeitungsprozesse für diese besonderen Zeugnisse auf.

**Abbildung 1**

Modell zur Beschreibung der Begegnung mit komplexen historischen Quellen



Das Modell identifiziert fünf Elemente der kognitiven Involviertheit: (1) **Aufmerksamkeitsfokus**: die Lernenden richten all ihre Aufmerksamkeit auf die Begegnung (z. B. Csikszentmihalyi & Robinson, 1990, 28); (2) **Imagination**: die Vergangenheit ist für die Lernenden bildlich vorstellbar (z. B. Green & Brock, 2000, 703); (3) **Perspektivennachvollzug**: die Lernenden verstehen die Perspektive des historischen Akteurs (z. B. Endacott & Brooks, 2013, 43); (4) **Kontextualisierung**: die Lernenden verknüpfen die Informationen mit ihrem Vorwissen (z. B. Huijgen et al., 2017, 112); (5) **(gefühlte) Erkenntnis**: die Lernenden glauben während der Begegnung, etwas über die Vergangenheit im Allgemeinen zu lernen (z. B. Busselle & Bilandzic, 2009, 330). Fünf Elemente machen laut des Modells die emotionale Involviertheit aus: (1) **Bewegt sein**: die Lernenden sind emotional ergriffen (z. B. Menninghaus et al., 2015, 12); (2) **persönliche Verbindung**: die Lernenden identifizieren sich mit dem Akteur (z. B. Brauer, 2016, 32); (3) **historische Unmittelbarkeit**: die Vergangenheit fühlt sich durch die Begegnung nah an (z. B. Lombard & Ditton, 1997, 7); (4) **Ehrfurcht**: die Lernenden empfinden eine tiefe Anerkennung (z. B. Cameron & Gatewood, 2003, 67); (5) **Irritation**: die Lernenden sind irritiert von den neuen Informationen (z. B. Brauer, 2019, 277). Drei Elemente setzen den

Rahmen der physischen Involviertheit: (1) Sinnliche Wahrnehmung: Der\*die Zeitzug\*in ist „zum Greifen“ nah (z. B. Drozdowski et al., 2016, 447); (2) Physische Interaktion: die Lernenden können Antworten auf ihre eigenen Fragen erhalten; (3) Physiologische Reaktion: die Lernenden zeigen körperliche Reaktionen, haben Tränen in den Augen oder Gänsehaut (z. B. Latham, 2013, 15).

Zuletzt beschreibt die „Einzigartigkeit des Moments“ ein Zusammenspiel kognitiver, emotionaler und physischer Involviertheit (Zachrich et al., 2020). Diese Elemente stellen mögliche Prozesse der Lernerfahrung dar, die nicht in jeder Begegnung und für jede Person von Bedeutung sind. Das „Aura-Erlebnis“ ist demnach durch hohe Ausprägungen auf mehreren dieser Subskalen gekennzeichnet. Wir legen dabei zum jetzigen Zeitpunkt keine bestimmte Zusammensetzung des „Aura-Erlebens“ über die Subskalen fest, weshalb wir im Folgenden in der Regel auf Basis der einzelnen Involviertheitsskalen argumentieren, nicht jedoch auf der Ebene des „Aura-Erlebens“.

### **Bedingungen der besonderen Lernerfahrung**

Welche Bedingungen erhöhen die Wahrscheinlichkeit, dass kognitive, emotionale und physische Involviertheit auftritt, und welche Konsequenzen hat sie? Wir haben hierzu das Modell von Zachrich et al. (2020) in ein Wirkungsmodell umstrukturiert und um Bedingungen und Konsequenzen erweitert (siehe Abbildung 2). Dabei unterscheiden wir bei den Bedingungen zwischen Merkmalen der Person und der Situation, denn theoretische wie empirische Arbeiten legen nahe, dass sowohl die Eigenschaften der Lernenden als auch die Merkmale der Situation einen Einfluss darauf haben, wie die Begegnung erlebt wird: In einer Studie von Green et al. (2012, 48) berichteten Frauen, stärker in eine Geschichte einzutauchen als Männer. Gnambs et al. (2014, 191) konnten zeigen, dass die Erfahrung von Transportation sowohl auf Personenmerkmale als auch auf situationale Merkmale zurückzuführen ist. Auch weil man davon ausgeht, dass die Unterrichtssituation einen Einfluss auf die Involviertheit in der Begegnung hat, empfehlen Lehrbücher zur Geschichtsdidaktik, den Zeitzugbesuch intensiv mit den Schüler\*innen vor- und nachzubereiten (z. B. Henke-Bockschatz, 2007, 354).

Die große Bedeutsamkeit der Begegnung schreiben Lehrkräfte und Schüler\*innen u.a. einer „Echtheit“ und „Authentizität“ des Zeitzugberichts zu. Empirische Untersuchungen in der Medienpsychologie bestätigten die Bedeutung der wahrgenommenen Echtheit (*perceived realism*) in Zusammenhang mit Erfahrungen wie Transportation (z. B. Green, 2004, 257), Immersion und Identifikation (z. B. Bilandzic & Busselle, 2011, 45) sowie emotionale Reaktionen der Rezipient\*innen (z. B. Pouliot & Cowen, 2007, 250). Einen Unterschied der

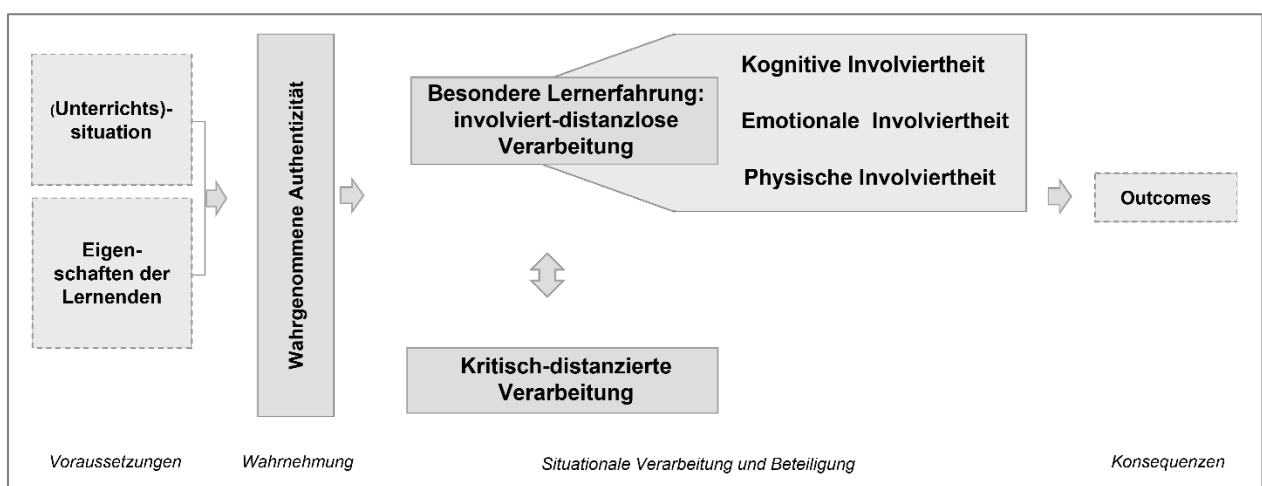
Erfahrung zwischen der Bezeichnung des Narratives als Fakt bzw. Fiktion zeigte sich allerdings in einer Reihe von Experimenten nicht konsistent (Green & Brock, 2000, 718). Mierwald et al. (2018, 292) konnten in einer Interventionsstudie wiederum Hinweise finden, dass authentisches Lernmaterial zum Thema „Einfluss der NS- Herrschaft auf das Zusammenleben von Juden und Nichtjuden“ einen positiven Einfluss auf epistemische Überzeugungen im Fach Geschichte hat. Allerdings wurden die vertonten Originalquellen von den Schüler\*innen als authentischer wahrgenommen als die ursprünglich schriftlich vorliegenden Quellen. Die genannten Studien implizieren, dass eine differenzierte Betrachtung von Authentizität notwendig ist. Ob die wahrgenommene Authentizität des Zeitzeugenberichts eine Rolle für die Lernerfahrung spielt, bleibt mit Blick auf die inkonsistente Befundlage eine empirische Frage. Wir haben deshalb im Wirkungsmodell die wahrgenommene Authentizität als Mediator zwischen Bedingung und Lernerfahrung berücksichtigt.

In Anlehnung an Vorderer (1992, 84) nehmen wir zudem an, dass es neben der starken Involviertheit mit dem Zeitzeugenbericht auch einen kritisch-distanzierten Verarbeitungsmodus gibt. Diesbezüglich geben bisherige Forschungsarbeiten Hinweise darauf, dass Transportation positiv mit Persuasion zusammenhängt (Green & Brock, 2000, 719) und Rezipient\*innen durch die Erfahrung des Eintauchens Inkonsistenzen in der Geschichte leichter akzeptieren, folglich weniger kritisch sind (Weber & Wirth, 2014, 136).

Die Studie von Bertram et al. (2017, 469) gab empirische Hinweise, dass die Begegnung mit Zeitzeug\*innen einen Einfluss auf beispielsweise die Motivation oder das Lernen der Schüler\*innen haben kann. Deshalb ergänzten wir das Modell schließlich um Konsequenzen der Lernerfahrung, die wir aber hier nicht untersuchten.

## Abbildung 2

Wirkmodell zur besonderen Lernerfahrung mit Zeitzeug\*innen



### **Fragestellung & Hypothesen**

Obens und Geißler (2010) stellten fest, dass es an Untersuchungen in „nahezu allen Aspekten dieses Bildungsarrangements“ (4) fehlt. Mit der vorliegenden Studie begegnen wir genau diesem Forschungsdesiderat und entwickelten als Grundlage für weitere empirische Studien ein standardisiertes Messinstrument zur Erfassung dieser besonderen Lernerfahrung. Dabei interessierte uns (i) was sich empirisch hinter der „Blackbox“ der Zeitzeugenbegegnung verbirgt und (ii) welche Prädiktoren auf Situations- und Personenebene die starke Involviertheit mit dem Zeitzeugenbericht vorhersagen können. Konkret orientierten wir uns dafür an den folgenden Fragen: (1) Lassen sich die verschiedenen theoretisch angenommenen Dimensionen der Erfahrung mit Zeitzeugenberichten empirisch trennen? (2) In welchem Zusammenhang stehen die erfassten Konstrukte mit der kritisch-distanzierten Verarbeitung? Wir erwarteten, dass eine starke Involviertheit negativ mit der kritisch-distanzierten Verarbeitung zusammenhängt. (3) Welche individuellen Faktoren der Schüler\*innen können eine hohe Involviertheit mit dem Zeitzeugenbericht vorhersagen? Wir nahmen an, dass Mädchen eine höhere Involviertheit berichten als Jungen; zudem sollten Schüler\*innen, die eine Frage gestellt haben, eine höhere Involviertheit berichten als Schüler\*innen, die nur zuhörten. (4) In der vierten Fragestellung interessierte uns, welche situationalen Merkmale hohe Involviertheit vorhersagen können. Mit Blick auf die Intensität der Vorbereitung hatten wir keine konkrete Erwartung, aber wir erwarteten, dass der zeitliche Abstand zwischen Begegnung und Studienteilnahme negativ mit der Involviertheit assoziiert ist. (5) Zuletzt interessierte uns, ob die Zusammenhänge zwischen den Bedingungen und der Begegnung durch die wahrgenommene Authentizität vermittelt werden.

### **Methode**

#### **Fragebogenentwicklung**

Der Prozess der Fragebogenentwicklung vollzog sich in mehreren Schritten. Aufbauend auf den theoretischen Überlegungen entwickelten wir in Fokusgruppen mit Expert\*innen aus der Bildungspraxis und -forschung, Geschichtsdidaktik und -wissenschaft standardisierte Items und passten Items aus validierten Fragebögen auf unseren Kontext an (z. B. Busselle & Bilandzic, 2009, 329–330). Der erste Entwurf umfasste 92 Items zu 17 Konstrukten der emotionalen, kognitiven und physischen Involviertheit, die auf einer 6-stufigen Rating-Skala beantwortet werden.

Die Pilotierung des Fragebogens fand mit  $N = 296$  Studierenden aus sieben Veranstaltungen der Universität Tübingen statt. Als Impuls sahen die Studierenden einen fünf-



minütigen Videoclip von Kurt Gutmann, der im Zweiten Weltkrieg ein Kind war und von seinen Erfahrungen während des Krieges und in der Nachkriegszeit berichtet.<sup>3</sup> Wir werteten die Daten mittels Reliabilitätsanalysen, konfirmatorischen und exploratorischen Faktorenanalysen aus. Wir kürzten den Fragebogen auf 60 Items für 14 Konstrukte. Der überarbeitete Fragebogen wurde in einem nächsten Schritt mit fünf Schüler\*innen in *Cognitive Labs* (Hilton, 2017) überprüft. Aufgrund von Missverständnissen bei wenigen Items passten wir Formulierungen an. Die Validierung des Fragebogens mit Schüler\*innen nach der Begegnung mit Live-Zeitzeug\*innen steht als letzter Schritt der Fragebogenentwicklung im Fokus der vorliegenden Studie und wird deshalb im Folgenden im Detail beschrieben.

### **Validierungsstudie**

Das Kultusministerium des Landes Baden-Württemberg hat die Studie und die Erhebung der Daten genehmigt (Datum der Genehmigung: 16. April 2019, Aktenzeichen: 31-6499.20/1274). Die Ethikkommission für psychologische Forschung der Universität Tübingen bestätigte, dass die Studie mit den ethischen Standards der Forschung am Menschen übereinstimmt (Datum der Genehmigung: 21. März 2019, Aktenzeichen: Trautwein\_2019\_02\_15\_149).

### *Forschungsdesign*

Unsere Studie wurde von zwei Institutionen unterstützt, die uns den Zugang zu Schüler\*innen nach einer Begegnung mit einem\*r Live-Zeitzeug\*innen ermöglichten. Deshalb enthält unsere Datengrundlage zwei unterschiedliche Settings von Zeitzeugenbegegnungen: (1) Das Koordinierende Zeitzeugenbüro ist ein Kooperationsprojekt der Bundesstiftung Aufarbeitung, der Stiftung Berliner Mauer, der Stiftung Gedenkstätte Berlin-Hohenschönhausen und der Beauftragten der Bundesregierung für Kultur und Medien. Das Büro vermittelt und finanziert DDR-Zeitzeug\*innen bundesweit an Schulen. Den Zeitzeug\*innen ist gemein, dass sie in der ehemaligen DDR aus verschiedenen Gründen inhaftiert waren. (2) Die Gedenkstätte Lindenstraße in Potsdam ist (unter anderem) ein ehemaliges Untersuchungsgefängnis des Ministeriums für Staatssicherheit. Die Gedenkstätte bietet individuelle Bildungsangebote für Schulklassen an, empfiehlt vor einem Gespräch mit einem\*r Zeitzeug\*in (ehemaliger Häftling) eine Führung durch das Gebäude. Zudem nutzten wir die Partnerschulen der Universitäten Konstanz und Tübingen, informierten über unser Projekt und riefen zur Teilnahme auf. Aufgrund terminlicher Schwierigkeiten im Schulalltag

---

<sup>3</sup> Das Zeitzeugenvideo wurde für eine Videoinstallation der Medienkünstler\*innen Ina Rommel und Stefan Kraus gedreht und für diese Studie zu einem kurzen Clip zusammengeschnitten: <https://vimeo.com/233487804>.

variierte der Erhebungszeitpunkt zwischen unmittelbar nach der Begegnung bis eine Woche später. Auch der Rahmen der Begegnung war nicht einheitlich: Einige Lehrkräfte luden den\*die Zeitzeug\*in in eine Klasse ein, während andere eine Großveranstaltung in der Aula organisierten, an der mehrere Klassen teilnahmen.

### *Stichprobe & Zeitzeug\*innen*

Insgesamt nahmen  $N = 668$  Schüler\*innen ( $M = 16,11$  Jahre,  $SD = 1,49$ ) aus 44 Lerngruppen der Klassenstufen 8 bis 13 an der Befragung teil. 52,1% der Befragten gaben an, weiblich zu sein, 45,8% männlich, 1,3% divers und 0,7% machten keine Angabe zu ihrem Geschlecht.<sup>4</sup> Bei 29 Gruppen lagen zwischen der Begegnung mit den Zeitzeug\*innen und der Erhebung weniger als 1 Tag, bei 8 Gruppen 1–2 Tage, bei 3 Gruppen 3–6 Tage und bei 4 Gruppen mehr als 6 Tage.

Die Daten wurden im Anschluss an Vorträge mit zehn Zeitzeug\*innen gesammelt, die von unterschiedlich vielen Schüler\*innen gesehen wurden. So haben 281 der 668 Schüler\*innen die Geschichte desselben Zeitzeugens gehört, während die Geschichte eines anderen Zeitzeugen nur vier Schüler\*innen gehört haben. Alle Zeitzeugenbesuche wurden als Vortrag mit anschließender Fragerunde organisiert.

### *Instrument*

Im Zentrum des Fragebogens stehen 50 Items, die 12 Konstrukte erfassen sollten. Die kognitive Involviertheit wurde über die Konstrukte Imagination, Aufmerksamkeitsfokus, Perspektivennachvollzug, (gefühlte) Erkenntnis und Kontextualisierung erhoben. Die emotionale Involviertheit wurde abgebildet durch die Konstrukte Bewegt sein, Identifikation, historische Unmittelbarkeit, Irritation und Einzigartigkeit des Moments. Die physische Involviertheit erfassten wir über die Konstrukte sinnliche Wahrnehmung und physiologische Reaktion.

Zur Erfassung der kritisch-distanzierten Verarbeitung adaptierten wir das von Bertram et al. (2014) entwickelte Instrument zur Erfassung der Einsicht in die Perspektivität von Zeitzeugenaussagen. Exploratorische Faktorenanalysen ergaben, dass lediglich drei der vierzehn Items hohe Ladungen auf dem zentralen Faktor und keine substantiellen Nebenladungen auf weitere Faktoren aufwiesen. Die folgenden drei Items bildeten folglich in

---

<sup>4</sup> Wenn wir im weiteren Verlauf der Studie von Mädchen und Jungen sprechen, beziehen wir uns auf jene Personen, die bei der Frage nach ihrem Geschlecht (weiblich, männlich, divers) das Kreuzchen bei Mädchen respektive Jungen gesetzt haben.

dieser Studie die kritisch-distanzierte Verarbeitung ab: „Zeitzeugen können uns wahre Auskünfte über die Vergangenheit geben, denn sie waren damals dabei“ (invers kodiert); „Viele Jahre später wissen Zeitzeugen noch genau, was früher alles passiert ist“ (invers kodiert); „Zeitzeugen vergessen möglicherweise, wie es genau gewesen ist.“ Cronbachs Alpha der reduzierten Skala war akzeptabel ( $\alpha = 0.68$ ).

Als Voraussetzungen der Lernerfahrung erhoben wir die Hintergrundvariablen Geschlecht und Interaktion zwischen Lernenden und Zeitzeug\*in („Hast du eine Frage gestellt?“). Zudem wurde in einem kurzen Lehrerfragebogen der Zeitraum zwischen Besuch und Befragung und eine Einschätzung zur Vorbereitung des Besuchs erfragt (1 = wenig vorbereitet bis 4 = intensiv vorbereitet).

Die wahrgenommene Authentizität erfassten wir über drei Items, um drei Facetten abzubilden: Authentizität des Berichts: „Der Bericht des Zeitzeugen war für mich sehr glaubwürdig“; Authentizität der Person: „Der Zeitzeuge hat sich uns gegenüber ehrlich geöffnet“; Authentizität der Quelle: „Die persönlichen Geschichten von jemandem, der erlebt hat, wovon er berichtet, haben die Vergangenheit lebendig gemacht.“ Die Items korrelierten im mittleren Bereich miteinander: Bericht und Person:  $r = .25$ ; Bericht und Quelle:  $r = .30$ ; Person und Quelle:  $r = .23$ .

### *Analyse*

Für unsere erste Forschungsfrage begannen wir in einem ersten Schritt mit exploratorischen Faktorenanalysen mit Einfachladungsstruktur und testeten dazu jede mögliche Itemkombination für ein Set an Faktoren und die Veränderung der Modellgüte bei Ausschluss einzelner Items aus dem Messmodell. Aufgrund der hohen Anzahl möglicher Itemkombinationen mussten sich diese Analysen jeweils auf einen Teilbereich (also eine Itemauswahl) beschränken. In einigen Fällen überprüften wir, ob eine Erweiterung der Faktorenanzahl zu einer Verbesserung der Modellgüte führte. Ziel war hier, die Konstrukte so zu spezifizieren, dass die Items keine Mehrfachladung zu anderen Konstrukten aufwiesen. Diese Berechnungen dienten als Grundlage für eine weitere Kürzung der Itemanzahl pro Konstrukt. In einem zweiten Schritt überprüften wir die vorgeschlagenen Modelle der jeweils mithilfe eindimensionaler Items erfassten Konstrukte mit konfirmatorischen Faktorenanalysen und berechneten Cronbach's Alpha zur Beurteilung der Homogenität der Skalen. Wir untersuchten zudem mittels zwei Intraklassenkorrelationskoeffizienten ( $ICC_{\text{unadjustiert}}$  und  $ICC_{\text{adjustiert}}$ ), ob die unterschiedlichen Zeitzeug\*innen im Sample für Unterschiede in der Involviertheit der Schüler\*innen verantwortlich sind. Hierzu verglichen wir die relativen

Varianzanteile der Skalen, die Unterschiede zwischen Klassen repräsentieren ( $ICC_{\text{unadjustiert}}$ ), mit solchen aus einem Modell, in dem wir für Unterschiede zwischen Zeitzeug\*innen (als Dummy-Variablen spezifiziert) kontrollierten ( $ICC_{\text{adjustiert}}$ ). Eine relativ kleine  $ICC_{\text{adjustiert}}$  würde dafür sprechen, dass die Unterschiede zwischen Klassen nicht (wesentlich) auf Unterschiede in der Schülerkomposition zurückgehen, sondern – bei deutlich höherer  $ICC_{\text{unadjustiert}}$  – auf Unterschiede seitens der Zeitzeug\*innen.

Zur Beantwortung der weiteren Forschungsfragen berechneten wir die latenten Korrelationen zwischen den Konstrukten und dem kritisch-distanzierten Faktor. Wir spezifizierten ein Regressionsmodell mit den persönlichen und situativen Merkmalen als Prädiktoren. Aus Gründen der Vergleichbarkeit der Modelle inkludierten wir alle hier im Fokus stehenden Variablen in jedem Modell als Korrelate, womit die Datengrundlage also jeweils identisch war. Zuletzt spezifizierten wir ein Mediationsmodell mit der wahrgenommenen Authentizität als Mediator. Wir verglichen, ob die standardisierten direkten Effekte aus dem Regressionsmodell nach Hinzunahme des Mediators kleiner wurden und prüften, ob die indirekten Effekte signifikant wurden. Die berichteten Regressionskoeffizienten wurden anhand der Varianz der abhängigen Variable sowie für kontinuierliche Prädiktoren zusätzlich an der Varianz der unabhängigen Variable standardisiert.

Wir bewerteten die Modellgüte nach dem *Root Mean Square Error of Approximation* (RMSEA), dem *Comparative Fit Index* (CFI) und dem *Standardized Root Mean Square Residual* (SRMR). Werte des RMSEA unter 0.05, CFI-Werte über 0.95 und SRMR-Werte unter 0.08 sprechen für eine akzeptable bis gute Modellpassung (Yu, 2002). Wir testeten die statistische Signifikanz jeweils auf dem 5%-Niveau. Die Analysen wurden mithilfe der Statistikprogramme Mplus 8.0 (Muthén & Muthén, 1998–2017) und SAS 9.4 (SAS Institute Inc., 2018) durchgeführt. Bei den konfirmatorischen Analysen in Mplus wurde die Clusterstruktur der Daten durch Korrektur der Standardfehler und Modelfit-Indizes über *Robust-Maximum-Likelihood*-Schätzung berücksichtigt.

## Ergebnisse

### Faktorenanalysen

Für die besondere Lernerfahrung standen 12 Konstrukte mit insgesamt 50 Items im Fokus dieses ersten Analyseschrittes. Nach der Prüfung, ob der Ausschluss von Items zu einer bedeutsamen Verbesserung der Modellpassung führte und Doppelladungen zu anderen Konstrukten aufwiesen, schlossen wir schrittweise 10 Items (20%) aus. Die

Interkorrelationsstruktur der Konstrukte für Sinnliche Wahrnehmung, Identifikation und Unmittelbarkeit wies einen Wert  $r > 1$  auf, weshalb wir diese drei Konstrukte in einer exploratorischen Faktorenanalyse gemeinsam untersuchten. Wir extrahierten aus den Ergebnissen einen Faktor mit vier Items, welcher das Eintauchen in die vergangene Zeit und das Spüren einer unmittelbaren Nähe zum\*r Zeitzeug\*in und zur Vergangenheit beschreibt und den wir als *historische Nähe* interpretieren. Die Items zum Konstrukt Irritation luden nicht eindeutig auf einen Faktor, weshalb diese ausgeschlossen wurden. Die Items des Konstrukts Imagination luden auf zwei Faktoren, weshalb wir die Items in die Faktoren *bildliche Imagination* und *abstrakte Vorstellungskraft* aufteilten. *Bildliche Imagination* beschreibt, dass Lernende lebendige Szenen zu dem Bericht im Kopf haben, während *abstrakte Imagination* erfasst, ob die Vergangenheit durch den Bericht für die Lernenden grundsätzlich vorstellbar war. Das Konstrukt der physiologischen Reaktion (mit zwei Items) schlossen wir aus, da die Items zu „stark“ für die Begegnung mit DDR-Zeitzeug\*innen formuliert waren (z. B. „Tränen in den Augen“). Denn diese Berichte waren häufig nicht so dramatisch, wie z. B. jene von Holocaust-Überlebenden. Folglich arbeiteten wir in den weiteren Untersuchungen nur mit den Dimensionen der kognitiven und emotionalen Involviertheit.

Die besondere Lernerfahrung mit den Zeitzeugenberichten zeichnet sich durch eine starke kognitive wie emotionale Involviertheit aus: Die kognitive Involviertheit lässt sich über die sechs Faktoren Aufmerksamkeitsfokus, bildliche Imagination, abstrakte Vorstellungskraft, Perspektivennachvollzug, gefühlte Erkenntnis und Kontextualisierung erfassen. Die emotionale Involviertheit mit dem Zeitzeugenbericht können wir über die drei Faktoren Bewegt sein, historische Nähe und Einzigartigkeit des Moments beschreiben. Das Modell mit diesen Konstrukten zeigt eine gute Modellpassung: RMSEA = 0.032, CFI = 0.968, SRMR = 0.039. Tabelle 1 gibt einen Überblick über die erfassten Konstrukte, ihre Kennwerte und Beispielitems. Ein Vergleich von  $ICC_{\text{unadjustiert}}$  und  $ICC_{\text{adjustiert}}$  zeigte, dass bei einer Kontrolle für Zeitzeug\*innen immer noch etwa die Hälfte der Varianz auf Klassenebene verbleibt und somit die unterschiedlichen Zeitzeug\*innen die Varianz der Involviertheit nur teilweise erklären können. Tabelle 2 zeigt die latenten Interkorrelationen zwischen den Faktoren. Mit Ausnahme von Kontextualisierung korrelierten alle Faktoren mittel bis stark miteinander. Die Interkorrelationen zeigten hierbei kein eindeutiges Muster mit Blick auf die Zuordnung zu emotionaler und kognitiver Involviertheit. Die Spezifikation von Second-Order-Faktoren führte zwar zu einer Verschlechterung der Modellpassung, wobei die Werte noch als akzeptabel betrachtet werden können: RMSEA = 0.039, CFI = 0.951, SRMR = 0.054.

**Tabelle 1**

Kennwerte der finalen Konstrukte zur Beschreibung der besonderen Lernerfahrung mit Zeitzeugenberichten.

Be- reich	Konstrukt	Item- anzahl	Beispielitem	<i>N</i>	<i>M</i>	<i>SD</i>	Cronbachs Alpha	ICC unadjustiert	ICC adjustiert	Range der std. Ladungen
Emotionale Involviertheit	Bewegt sein	3	Der Bericht des ZZ ging mir nahe.	667	4.01	1.04	0.87	0.20	0.10	0.78–0.87
	Historische Nähe	4	Obwohl der ZZ von früher erzählt hat, habe ich eine Nähe zu dieser Zeit gespürt.	667	3.79	0.89	0.78	0.17	0.07	0.62–0.72
	Einzigkeit	4	Die Begegnung mit dem ZZ-Bericht war für mich einzigartig.	668	4.03	0.98	0.86	0.15	0.11	0.65–0.87
Kognitive Involviertheit	Aufmerksam- keitsfokus	3	Ich war bei dem ZZ-Bericht voll und ganz bei der Sache.	667	4.46	0.98	0.86	0.21	0.12	0.71–0.89
	Bildliche Imagination	3	Als ich dem ZZ zugehört habe, sind mir Szenen zu dem Bericht in den Kopf gekommen.	668	4.39	1.03	0.85	0.09	0.04	0.75–0.90
	Abstrakte Vorstellung	3	Für mich war die Zeit, von der der ZZ berichtet hat, sehr gut vorstellbar.	668	4.68	0.73	0.76	0.10	0.04	0.64–0.76
	Perspektiven- nachvollzug	4	Ich kann die Empfindungen des ZZ verstehen.	668	4.86	0.72	0.82	0.13	0.06	0.69–0.78
	Gefühlte Erkenntnis	5	Durch den ZZ-Bericht habe ich verstanden, wie es damals war.	666	4.66	0.71	0.82	0.07	0.02	0.46–0.78
	Kontextuali- sierung	3	Ich habe die Begegnung mit meinem Wissen über die Zeit in Verbindung gebracht.	668	4.80	0.88	0.88	0.05	0.04	0.71–0.80

*Notiz.* ICC<sub>unadjustiert</sub> = relative Varianzanteile der Skalen, die Unterschiede zwischen Klassen repräsentieren; ICC<sub>adjustiert</sub> = relative Varianzanteile der Skalen, wenn für Unterschiede zwischen Zeitzeug\*innen kontrolliert wurde. Grundlage für den Range der standardisierten Ladungen ist das konfirmatorische Faktorenmodell mit eindimensionalen Faktoren.

**Tabelle 2**

Latente Interkorrelationen zwischen den Faktoren emotionaler und kognitiver Involviertheit.

	1	2	3	4	5	6	7	8
1 Bewegt sein	1							
2 Historische Nähe	.75	1						
3 Einzigartigkeit	.76	.69	1					
4 Aufmerksamkeitsfokus	.55	.58	.61	1				
5 Bildliche Imagination	.54	.60	.50	.45	1			
6 Abstrakte Vorstellung	.51	.72	.54	.55	.70	1		
7 Perspektivennachvollzug	.58	.57	.53	.46	.55	.72	1	
8 Gefühlte Erkenntnis	.50	.57	.52	.38	.54	.78	.78	1
9 Kontextualisierung	.21	.21	.26	.27	.27	.31	.37	.30

Notiz.  $N = 668$  Schüler\*innen.

### Zusammenhangsanalysen

#### *Überwältigt vs. kritisch-distanziert*

In der zweiten Fragestellung interessierte uns, wie die besondere Lernerfahrung mit der kritisch-distanzierten Verarbeitung zusammenhängt. Die Korrelationen bestätigten unsere Erwartungen: Die Involviertheit-Faktoren hingen statistisch signifikant negativ mit der kritisch-distanzierten Verarbeitung zusammen. Der Range lag im Bereich von  $r = -.16$ ,  $SE = 0.07$  mit Kontextualisierung bis  $r = -.58$ ,  $SE = 0.04$  mit gefühlter Erkenntnis. Das Modell zeigte gute Fit-Indizes:  $RMSEA = 0.032$ ,  $CFI = 0.964$ ,  $SRMR = 0.041$ .

#### *Einfluss von Personen- und Situationsmerkmalen*

In der dritten und vierten Fragestellung untersuchten wir, welche individuellen Faktoren der Schüler\*innen und Situationsmerkmale eine starke Involviertheit vorhersagten. Hierfür spezifizierten wir ein Modell mit den Prädiktoren Geschlecht (dummy-kodiert: 0 = Mädchen, 1 = Jungen) und Interaktion mit dem\*der Zeitzeug\*in, zeitlicher Abstand zwischen Begegnung und Erhebungstermin und der Intensität der Vorbereitung des Zeitzeugenbesuchs. Mädchen gaben konsistent an, stärker involviert zu sein als Jungen: Der stärkste Unterschied zeigte sich mit dem Faktor bewegt sein:  $b = -0.29$ ,  $SE = 0.04$ ,  $p < .001$ . Der kleinste Unterschied lag zwischen Geschlecht und bildlicher Imagination:  $b = -0.11$ ,  $SE = 0.04$ ,  $p = .012$ . Bei dem Faktor Kontextualisierung zeigte sich kein statistisch signifikanter Unterschied. Jungen hingegen verarbeiteten die Informationen aus dem Zeitzeugenbericht stärker kritisch-distanziert als

Mädchen ( $b = 0.17$ ,  $SE = 0.06$ ,  $p = .003$ ). Wenn Schüler\*innen dem\*r Zeitzeug\*in eine Frage stellten, zeigten sie auf allen Konstrukten eine statistisch signifikant stärkere Involviertheit mit dem Bericht. Die Interaktion mit dem\*r Zeitzeugen\*in führte zu einer Reduktion der kritisch-distanzierten Verarbeitung ( $b = -0.20$ ,  $SE = 0.05$ ,  $p < .001$ ). Für die Prädiktoren zeitlicher Abstand und Vorbereitung des Besuchs zeigten sich keine statistisch signifikanten Zusammenhänge. Der Determinationskoeffizient der latenten Variablen reichte von .02 (Kontextualisierung)  $\leq R^2 \leq .15$  (Bewegt sein). Das Regressionsmodell wies gute Modellfit-Werte auf: RMSEA = 0.033, CFI = 0.956, SRMR = 0.038. Die Ergebnisse bestätigten unsere Erwartungen mit Blick auf die Bedeutung persönlicher Merkmale in der Lernerfahrung. Für die Bedeutung der Intensität der Vorbereitung und den zeitlichen Abstand fanden wir kaum Hinweise. Das ist möglicherweise auf fehlende Varianz zurück zu führen (z. B. wurde kein Besuch „stark“ vorbereitet).

#### *Wahrgenommene Authentizität als Mediator*

In der letzten Fragestellung interessierte uns, ob die wahrgenommene Authentizität die Effekte der Prädiktoren auf die Involviertheit mediiert. Hierfür ergänzten wir das Regressionsmodell um drei manifeste Items zur Erfassung der drei Facetten von Authentizität, verglichen die Unterschiede zwischen den totalen Effekten ( $c'$ , aus einem Regressionsmodell ohne Moderatoren) und den direkten Effekten und testeten die indirekten Effekte (z. B. Effekt von Geschlecht auf Bewegt sein mediiert über die Authentizität des Berichts) auf Signifikanz.<sup>5</sup> Hierfür testeten wir, ob die Pfade von Prädiktor auf Mediator und von Mediator auf die abhängige Variable signifikant waren. Wir fanden das folgende Muster: Die direkten Effekte von Geschlecht auf Involviertheit wurden kleiner bzw. waren nach Hinzunahme der Mediatoren nicht mehr statistisch signifikant. Die kleinste Differenz fanden wir beim Faktor Kontextualisierung ( $c' - c = -0.08$ ), die größte Differenz zeigte sich beim Faktor Bewegt sein ( $c' - c = -0.18$ ). Die indirekten Effekte von Geschlecht über die Authentizität des Berichts sowie über die Authentizität der Quelle auf alle Involviertheitskalen waren statistisch signifikant – nicht jedoch über die Authentizität der Person (Ehrlichkeit). Auch der Effekt der Interaktion auf alle Involviertheitsfaktoren verringerte sich nach Hinzunahme der Mediatoren, jedoch nur marginal ( $c' - c \leq 0.05$ ).

Die Effekte der Mediatoren ( $b$ ) auf die Involviertheitsfaktoren waren insbesondere für die Facetten Authentizität des Berichts und der Quelle statistisch signifikant. Der stärkste Effekt

---

<sup>5</sup> Wir verwenden hier die typischen Begriffe von Mediationsmodellen. Allerdings implizieren wir aufgrund des Studiendesigns keine kausalen Effekte.



zeigte sich von Glaubwürdigkeit des Berichts auf Perspektivennachvollzug ( $b = 0.45$ ,  $SE = 0.05$ ). Den kleinsten Effekt fanden wir von Glaubwürdigkeit auf Kontextualisierung ( $b = 0.13$ ,  $SE = 0.07$ ). Dieser war als einziger jedoch nicht statistisch signifikant. Mit Blick auf die Authentizität der Quelle fanden wir den stärksten Effekt auf den Faktor historische Nähe ( $b = 0.51$ ,  $SE = 0.04$ ) und den schwächsten Effekt auf den Faktor Kontextualisierung ( $b = 0.17$ ,  $SE = 0.04$ ). Die Authentizität der Person zeigte lediglich statistisch signifikante Zusammenhänge zu Kontextualisierung ( $b = 0.12$ ,  $SE = 0.04$ ), gefühlter Erkenntnis ( $b = 0.12$ ,  $SE = 0.04$ ) und kritisch distanzierter Verarbeitung ( $b = 0.14$ ,  $SE = 0.06$ ).

Ein Vergleich der Determinationskoeffizienten in Regressions- und Mediationsmodell zeigte, dass durch Hinzunahme der Mediatoren deutlich mehr Varianz der Involviertheitsfaktoren erklärt werden konnte:  $.09$  (Kontextualisierung)  $\leq \Delta R^2 \leq .48$  (Perspektivennachvollzug). Das Mediationsmodell wies gute Modellfit-Werte auf: RMSEA = 0.032, CFI = 0.958, SRMR = 0.039. Die Ergebnisse der Mediationsanalysen geben Hinweise, dass die wahrgenommene Authentizität, insbesondere gemessen über die Glaubwürdigkeit des Berichts und die allgemeine Zustimmung zur Bedeutsamkeit der Quelle, eine besondere Rolle für die Verarbeitung des Berichts spielen und die persönlichen Merkmale nicht unmittelbar auf die Verarbeitung wirken.

### **Diskussion, Limitationen und Ausblick**

Ziele der vorliegenden Studie umfassten (1) die Entwicklung eines standardisierten Messinstruments, das es ermöglicht, die Lernerfahrung mit Zeitzeug\*innen besser zu verstehen, (2) die Analyse, welche Zusammenhänge zwischen persönlichen und situativen Merkmalen und der Lernerfahrung sowie zwischen der wahrgenommenen Authentizität und der Lernerfahrung bestehen. Das von uns zugrunde gelegte theoretische Modell erwies sich als produktiv, weil tatsächlich viele Faktoren auch empirisch bestätigt werden konnten, auch wenn die finale Faktorstruktur teilweise abwich von den postulierten Faktoren. Insbesondere konnten wir die physische Involviertheit mit dem Fragebogen nicht reliabel erfassen. Zudem fanden wir zwar bestätigende Hinweise auf die Zuordnung der Dimensionen zu den Faktoren der kognitiven und emotionalen Involviertheit, allerdings waren diese weniger eindeutig als erwartet. Der standardisierte Fragebogen zur Erfassung der emotionalen und kognitiven Dimension des „Aura-Erlebnisses“ in der Begegnung mit Zeitzeug\*innen kann nun in weiteren Studien eingesetzt werden, auch wenn eine Weiterentwicklung und erneute Prüfung des Instruments sinnvoll erscheinen, um möglicherweise weitere Teil-Prozesse der Lernerfahrung identifizieren zu können.

Die vielleicht größte Einschränkung der vorliegenden Studie findet sich im Studiendesign: Die Thematik und das Setting der Zeitzeugenbegegnung variierten kaum, was letzten Endes zu einer eingeschränkten Varianz geführt haben könnte, die typischerweise dazu führt, dass die Anzahl und Ausprägtheit von Sub-Dimensionen unterschätzt wird. Dies wiederum könnte dazu führen, dass die Ergebnisse der Studie nicht direkt auf andere Thematiken und das Setting übertragbar wären. Von daher wäre es wichtig zu untersuchen, ob beispielsweise die physische Involviertheit in einer Studie mit stärker variierendem Setting und mit anderen Messmethoden (z. B. Hautwiderstand) reliabler erfasst werden kann.

Für weitergehende Forschung bietet sich auch eine explizite Verknüpfung mit der modernen Forschung zur Unterrichtsqualität an (vgl. Kunter & Trautwein, 2013). Zeitzeugenberichte aktivieren die Lernenden kognitiv und emotional. Damit aber die Begegnung nicht mit einem Verlust an kritischer Distanz, sondern einem Zuwachs an historischem Denken einhergeht, braucht es eine Einbettung in ein kognitiv aktivierendes Unterrichtssetting auf einer Meta-Ebene sowie die konstruktive Unterstützung durch die Lehrkraft, damit Schüler\*innen nicht mit ihren Emotionen allein gelassen werden. Trautwein, Schreiber und Hasenbein (2021) wiesen in ihrem Positionspapier zur Fachspezifität von Unterrichtsqualitätsmerkmalen explizit auf die besondere Bedeutung der konstruktiven Unterstützung im Geschichtsunterricht hin. Damit das Unterrichtssetting aber nicht nur zu einer „Kultivierung der Affekte“ (Mitscherlich, 1983, zitiert nach von Borries, 1994, 90), sondern auch zu einem tiefgreifenden Verständnis über die Veränderbarkeit von Erinnerung und die Perspektivität von Zeitzeugenaussagen beiträgt, ist eine kognitiv aktivierende Rahmung der Begegnung notwendig. Nur wenn wir die Prozesse dieser besonderen Lernerfahrung in Gänze verstehen, können wir daraus eine angemessene Methodik und Gestaltung für den Unterricht ableiten, welche das „Aura-Erlebnis“ mit Zeitzeug\*innen zu einem „Katalysator“ für das Lernen im Geschichtsunterricht machen.

---

**Literatur**

- Angvik, M., & von Borries, B. (1997). *Youth and history. A comparative European survey on historical consciousness and political attitudes among adolescents*. Körber-Stiftung.
- Bertram, C. (2020). Live oder medial vermittelt: Was lernen Schülerinnen und Schüler mit Zeitzeuginnen und Zeitzeugen? In A.-B. Rothstein & S. Pilzweiger-Steiner (Hrsg.), *Entgrenzte Erinnerung* (S. 247–266). De Gruyter.
- Bertram, C., Wagner, W. & Trautwein, U. (2014). Zeitzeugenbefragungen im Geschichtsunterricht: Entwicklung eines Kurzinstruments für die Wirksamkeitsmessung. In T. Arand & M. Seidenfuß (Hrsg.), *Neue Wege - neue Themen - neue Methoden? Ein Querschnitt aus der geschichtsdidaktischen Forschung des wissenschaftlichen Nachwuchses* (S. 191–208). V&R Academic.
- Bertram, C., Wagner, W., & Trautwein, U. (2017). Learning historical thinking with oral history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54(3), 444–484. <https://doi.org/10.3102/0002831217694833>
- Bilandzic, H., & Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, 36(1), 29–50. <https://doi.org/10.1515/COMM.2011.002>
- Brauer, J. (2016). “Heiße Geschichte”? Emotionen und historisches Lernen in Museen und Gedenkstätten. In S. Willner, G. Koch, & S. Samida (Hrsg.), *Doing History - Performative Praktiken in der Geschichtskultur* (S. 29–44). Waxmann.
- Brauer, J. (2019). Gefühlte Geschichte? Emotionen, Geschichte und historisches Lernen. *Vierteljahresschrift Für Wissenschaftliche Pädagogik*, 95, 272–283. <https://doi.org/10.30965/25890581-09501022>
- Brauer, J., & Lücke, M. (Hrsg.) (2013). *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven*. V&R unipress.
- Brüning, C. I. (2018). *Holocaust Education in der heterogenen Gesellschaft*. Wochenschau.
- Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, 12(4), 321–347. <https://doi.org/10.1080/15213260903287259>
- Cameron, C. M., & Gatewood, J. B. (2003). Seeking numinous experiences in the unremembered past. *Ethnology*, 42(1), 55–71.
- Csikszentmihalyi, M., & Robison, R. E. (1990). *The art of seeing: An interpretation of the aesthetic encounter*. Getty Center for Education in the Arts.
- Drozdowski, D., De Nardi, S., & Waterton, E. (2016). Geographies of memory, place and

- identity: Intersections in remembering war and conflict. *Geography Compass*, 10(11), 447–456. <https://doi.org/10.1111/gec3.12296>
- Endacott, J., & Brooks, S. (2013). An updated theoretical and practical model for promoting historical empathy. *Social Studies Research and Practice*, 8(1), 41–58. <http://socstrp.org/>
- Gnambs, T., Appel, M., Schreiner, C., Richter, T., & Isberner, M. B. (2014). Experiencing narrative worlds: A latent state-trait analysis. *Personality and Individual Differences*, 69, 187–192. <https://doi.org/10.1016/j.paid.2014.05.034>
- Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, 38(2), 247–266. <https://doi.org/10.1207/s15326950dp3802>
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Green, M. C., Chatham, C., & Sestir, M. A. (2012). Emotion and transportation into fact and fiction. *Scientific Study of Literature*, 2(1), 37–59. <https://doi.org/10.1075/ssol.2.1.03gre>
- Henke-Bockschatz, G. (2007). Zeitzeugenbefragung. In U. Mayer, H.-J. Pandel, & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 354–369). Wochenschau.
- Huijgen, T., van Boxtel, C., van de Grift, W., & Holthuis, P. (2017). Toward historical perspective taking: Students' reasoning when contextualizing the actions of people in the past. *Theory and Research in Social Education*, 45(1), 110–144. <https://doi.org/10.1080/00933104.2016.1208597>
- Klika, D. (2018). Bildung und Emotion - Historisch-systematische Zugänge. In M. Huber & S. Krause (Hrsg.), *Bildung und Emotion* (S. 75–90). Springer.
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. UTB.
- Lanman, B. A., & Wendling, L. M. (2006). *Preparing the next generation of oral historians*. Altamira.
- Latham, K. F. (2013). Numinous experiences with museum objects. *Visitor Studies*, 16(1), 3–20. <https://doi.org/10.1080/10645578.2013.767728>
- Lombard, M., & Ditton, T. (1997). At the heart of it all: the concept of presence. *Journal of Computer-Mediated Communication*, 3(2), 1–36.
- Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PLoS ONE*, 10(6), 1–33. <https://doi.org/10.1371/journal.pone.0128451>

- Mierwald, M., Lehmann, T., & Brauch, N. (2018). Changing epistemological beliefs in student labs: Authentic learning materials as a chance to foster the development of academically adequate beliefs in the domain of history? *Unterrichtswissenschaft*, 46(3), 279–297. <https://doi.org/10.1007/s42010-018-0019-7>
- Obens, K., & Geißler-Jagodzinski, C. (2010). „Dann sind wir ja auch die letzte Generation, die davon profitieren kann“. *Erste Ergebnisse einer empirischen Mikrostudie zur Rezeption von Zeitzeugengesprächen bei Jugendlichen/jungen Erwachsenen*. [https://www.vielfalt-mediathek.de/wp-content/uploads/2020/12/forschungsbericht\\_zeitzeugengesprache.pdf](https://www.vielfalt-mediathek.de/wp-content/uploads/2020/12/forschungsbericht_zeitzeugengesprache.pdf)
- Pandel, H.-J. (2012). *Geschichtsdidaktik. Eine Theorie für die Praxis*. Wochenschau.
- Pouliot, L., & Cowen, P. S. (2007). Does perceived realism really matter in media effects? *Media Psychology*, 9(2), 241–259. <https://doi.org/10.1080/15213260701285819>
- Sabrow, M. (2012). Der Zeitzeuge als Wanderer zwischen den Welten. In M. Sabrow & N. Frei (Hrsg.), *Die Geburt des Zeitzeugen nach 1945* (S. 13–32). Wallstein.
- Schreiber, W. (2002). Reflektiertes und (selbst-)reflexives Geschichtsbewusstsein durch Geschichtsunterricht fördern - ein vielschichtiges Forschungsfeld der Geschichtsdidaktik. In B. Schönemann, W. Schreiber, & H. Voit (Hrsg.), *Zeitschrift für Geschichtsdidaktik. Grundfragen, Forschungsergebnisse, Perspektiven* (S. 18–34). Wochenschau.
- Schreiber, W., Körber, A., von Borries, B., Krammer, R., Leutner-Ramme, S., Mebus, S., Schöner, A., & Ziegler, B. (2006). *Historisches Denken. Ein Kompetenz-Strukturmodell*. ars una.
- Trautwein, U., Schreiber, W., & Hasenbein, L. (2021). Geschichtsdidaktische Forschung und allgemeine Unterrichtsforschung: Unbekannte Verwandte? *Unterrichtswissenschaft*, 49, 261–272. <https://doi.org/10.1007/s42010-021-00120-z>
- Uhl, H. (2012). Orte und Lebenszeugnisse - “Authentizität” als Schlüsselkonzept in der Vermittlung der NS-Verfolgungs- und Vernichtungspolitik. In M. Rössner & H. Uhl (Hrsg.), *Renaissance der Authentizität? Über die neue Sehnsucht nach dem Ursprünglichen* (S. 258–284). transcript.
- von Borries, B. (1994). Von gesinnungsbildenden Erlebnissen zur Kultivierung der Affekte? Über die Ziele und Wirkungen von Geschichtslernen. *Studien Zur Internationalen Schulbuchforschung*, 76, 67–92.
- von Plato, A. (2001). Chancen und Gefahren des Einsatzes von Zeitzeugen im Unterricht. *BIOS Zeitschrift Für Biographieforschung, Oral History Und Lebensverlaufsanalysen*,

14(2), 134–139.

Vorderer, P. (1992). *Fernsehen als Handlung. Fernsehfilmrezeption aus motivationspsychologischer Perspektive*. Edition Sigma.

Weber, P., & Wirth, W. (2014). When and how narratives persuade: The role of suspension of disbelief in didactic versus hedonic processing of a candidate film. *Journal of Communication*, 64(1), 125–144. <https://doi.org/10.1111/jcom.12068>

Wierling, D. (1997). Oral History. In K. Bergmann, K. Fröhlich, A. Kuhn, J. Rüsen, & G. Schneider (Hrsg.), *Handbuch der Geschichtsdidaktik* (S. 236–239). Kallmeyer.

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (30th ed.). University of California.

Zachrich, L., Weller, A., Baron, C., & Bertram, C. (2020). Historical experiences: A framework for encountering complex historical sources. *History Education Research Journal*, 17(2), 243–275.

---

# 5

## STUDIE 3

---

Zachrich, L., Wagner, W., Bertram, B. & Trautwein, U. (2022). Really? It depends! How authentic learning material affects involvement and learning with personal stories of the past. *Manuskript eingereicht zur Publikation.*

**Abstract**

Authentic learning is believed to possess qualities that boost student motivation and learning. However, prior studies have not consistently found the assumed positive effects of authentic learning, thus highlighting the need to better understand the underlying learning processes. Here, we examined effects of authentic learning material on learners' involvement and different facets of perceived authenticity of the material. We conducted an experimental online study with 336 adult participants in three conditions varying in the framing of their authenticity. We presented videos in which people talked about experiences in war and daily life in the past and labeled these clips eyewitness accounts, eyewitness accounts played by actors, or reconstructed accounts played by actors. Results revealed that treatment condition particularly affected participants' emotional involvement with the story and that especially one facet of perceived authenticity (Vividness of History) predicted learners' involvement.

*Keywords:* authentic learning, history education, learner's involvement, perceived authenticity



## 1. Introduction

Authentic learning is characterized by its relationship with the *real world*, and this real-world character is believed to make authentic learning relevant, motivating, and particularly successful (Fougt et al., 2019; Rule, 2006). Authentic learning uses a wide variety of (simulated) real-world materials and problems (Betz et al., 2016; Herrington & Oliver, 2000). For example, students come to science labs and learn how to be a physicist rather than just learning about physics in a classroom (Lombardi, 2007). Similarly, students visit a church to gain, among others, knowledge about objects found in churches as well as their functions (Riegel & Kindermann, 2016). In history education, students work with primary documents and sources referring to the Pledge of Allegiance (Rule, 2006), or people who witnessed a certain historical event are shown on videos or visit classrooms in person.

In recent decades, the concept of authentic learning has received an enthusiastic response in educational practice and has triggered a substantial body of empirical work (e.g., Fougt et al., 2019; Herrington & Oliver, 2000; Shaffer & Resnick, 1999). Most researchers agree that authentic learning offers a promising instructional approach for fostering students' motivation and achievement (Betz et al., 2016; Fougt et al., 2019; Herrington & Oliver, 2000; Shaffer & Resnick, 1999).

However, empirical research has drawn a somewhat inconsistent picture of the effects of authentic learning on several learning outcomes (e.g., Itzek-Greulich et al., 2015; Nachtigall & Rummel, 2021). Not all studies found positive effects on the major outcome variables and, even more troubling, there have been reports (e.g., Bertram et al., 2017) that some forms of authenticity might even interfere with some learning goals. Authenticity clearly matters in some way or another, but the inconsistency of results highlights the need to better understand how exactly authentic learning material affects the learning process. Consequently, we investigated the effects of authentic learning material (eyewitness reports of the past) on perceived authenticity, processing, and involvement with the stories and learning outcomes in an experimental setting.

### 1.1 Authentic Learning

The concept of authentic learning is rooted in situated learning theory, which postulates that students acquire transferable knowledge best in situations that represent the way it will be used in real life (Herrington & Oliver, 2000). Studies on authentic learning are based on various meanings of authenticity in learning contexts (Herrington & Oliver, 2000; Shaffer & Resnick, 1999), for example, real-world authenticity referring to the real-world character of a learning

situation, or personal authenticity, referring to the honesty of the teacher or student (e.g., Foug et al., 2019). Betz et al. (2016) identified a common notion of authentic learning in several studies: Authenticity refers to the approximation of learning methods, materials, and situations to the methods, materials, and situations in people's everyday and professional lives. Scientists conceptualized authenticity as a continuum that ranges from low to high authenticity, depending on the real-life context (Betz et al., 2016).

The concept of authentic learning is discussed in various subject areas (Yıldız, 2021), but independently of the discipline, authentic learning is considered to boost students' motivation and achievement: the more authentic, the more promising the learning environment (e.g., Betz et al., 2016; Cronin, 1993; Herrington & Oliver, 2000). However, previous empirical research has drawn an inconsistent picture of the effects of authentic learning on students' learning outcomes. Some studies found positive effects of authentic learning on students' interest and motivation (e.g., Bertram et al., 2017; Itzek-Greulich et al., 2017), but others did not (e.g., Feddermann et al., 2021; Nachtigall & Rummel, 2021). These inconsistencies can also be found for the effects of authentic learning on students' achievement (e.g., Bertram et al., 2017; Itzek-Greulich et al., 2015; Nachtigall & Rummel, 2021; Riegel & Kindermann, 2016).

The mixed empirical support for authentic learning raises questions: What might explain the inconsistencies between the theoretical assumptions and empirical results? A promising avenue for answering this question involves taking a closer look at what exactly happens when someone is confronted with authentic learning material. In other words: How is the *learning process* affected by authenticity?

## **1.2 Situational Processing and Involvement in Authentic Learning Settings**

The present research is based on a model (see Figure 1) that provides a testable description of the underlying learning processes in authentic learning settings. To this end, we adapted Betz et al.'s (2016) model, which Nachtigall et al. (2019) simplified, and added a component called *situational processing and involvement*.

Betz et al. (2016) described a process of authentication that occurs through the interaction of personal characteristics and characteristics of the learning setting and results in an individual feeling of authenticity (see Box 3 in Figure 1). This subjective perception is determined by the learner's attribution of authenticity to the learning setting. Perceived authenticity is expected to mediate the effects of the authentic learning setting on several learning outcomes. However, the few available studies that have investigated this mediating

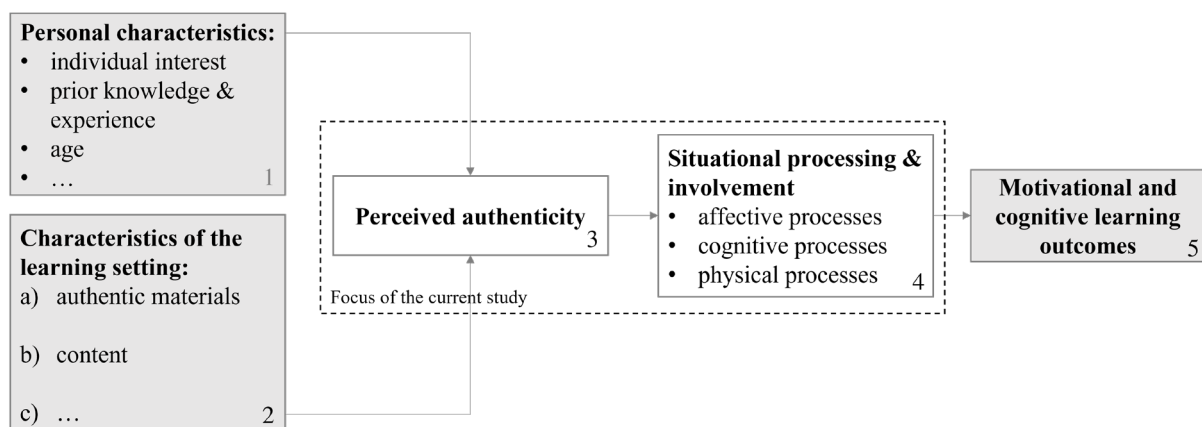
role of perceived authenticity on learning outcomes did not consistently find the expected relationships (e.g., Betz, 2018; Nachtigall & Rummel, 2021; Stamer et al., 2021).

In the past, stated simply, research on authentic learning examined whether authentic learning “automatically” leads to good learning outcomes (e.g., Feddermann et al., 2021; Itzek-Greulich et al., 2015; Riegel & Kindermann, 2016), whereas the mental processes involved were largely ignored (Schwan et al., 2017). However, it might be important to differentiate between several different mental processes (Bilandzic & Busselle, 2011; Zachrich et al., 2020) that come along with authentic learning and that may differentially affect the learning outcomes (see Box 4 in Figure 1).

More specifically, through their real-life connection, authentic learning materials might trigger cognitive processes (e.g., being more attentive than with regular learning material; e.g., Csikszentmihalyi, 2014), affective processes (e.g., feeling a personal connection to the learning content; e.g., Green & Brock, 2000), and even physical processes when leaving the classroom to visit, for example, historic sites (e.g., Latham, 2013). Importantly, some of these processes might not always help students reach certain learning goals. For instance, affective processes might impede learners’ ability to transfer their knowledge to a more abstract level independent of the authentic learning situation. In conclusion, opening the black box of the learning process in authentic learning situations is needed to better understand the effects of authentic learning.

### Figure 1

*Model of Perceived Authenticity and Situational Processing—an Extended Version of the Betz et al. (2016) Model*



### 1.3 Learning with Personal Stories from the Past

The divergent learning processes that might occur with authentic learning materials are of key interest in history education where authentic learning plays an important yet sometimes

contested role. Here, personal stories of the past are highly authentic learning materials (Ballis & Schwendemann, 2021; De Jong, 2018) that teachers typically use with one of two traditions that might trigger different learning processes: On the one hand, (a) eyewitnesses of the past who suffered injustices or violence give witness to the past and are supposed to teach students lessons about the difficult chapters of history (Sabrow, 2012). On the other hand, (b) in the Oral History tradition, a methodological approach that was established in science of history, personal stories about the past offer insights into daily struggles in life and give “small people” a voice in historiography (as opposed to writing history through the voices of leaders). With these kinds of stories, teachers want to foster students’ abilities to critically handle the materials (Henke-Bockschatz, 2014). In both traditions, students learn about the past through people who witnessed a certain historical event, time, or place. Students encounter the story by actually meeting the person, watching videos, reading biographical documents (Bertram et al., 2017), or working on oral history projects (Whitman, 2004).

Although students appreciate learning from personal accounts (e.g., Ballis & Schwendemann, 2021; Brüning, 2018), they are not necessarily able to critically reflect on the story—which is one learning goal of history lessons in school in Western democracies. Indeed, in-person and digital accounts have the power to overwhelm students with emotional content and an immediate connection to the past—students feel they are learning “the truth,” not “a truth” (Ballis & Schwendemann, 2021; Bertram et al., 2017; Brüning, 2018). In line with research on the inconsistent effects of authentic learning in various subjects, learning about the past from personal stories might not always lead to the desired learning outcomes (i.e., critically reflecting on the story, orientation) but might even backfire. However, empirical studies on the effects of authentic learning material on these different mental processes are lacking.

#### **1.4 The Current Study**

We argue that students’ involvement plays a crucial role in the chain of effects of authentic learning on learning outcomes and that this involvement includes a set of different processes, some that might hinder learning outcomes and some that might foster them. We examined the effects of authentic learning material on learners’ involvement with the material and perceived authenticity. We focused on two historical topics to cover the wide range of topics in which these materials are used: (a) daily experiences around the fall of the Berlin Wall, and (b) experiences of people who were children during World War II. We conducted an online experiment in which we varied only the introduction to the video about firsthand experiences and thereby designed three authenticity conditions—videos showing eyewitness accounts that were labeled (a) original eyewitness accounts, (b) eyewitness accounts played by actors, or (c)

reconstructed accounts played by actors—that we argue contribute to authentic learning with these materials. We were guided by the following three research questions.

1. What effects do the authenticity of personal stories about the past (see Box 2 in Figure 1) have on situational processing and cognitive, affective, and physical involvement with the account (Box 4)?
2. What effects do the authenticity of personal stories about the past (Box 2) have on (a) situational interest in history, (b) sense of insight, (c) critical distance, and (d) remembering the information (Box 5)?
3. Does perceived authenticity (Box 3) mediate the effect of authenticity (Box 2) on processing and involvement, situational interest, sense of insight, and critical distance (Boxes 4 and 5)?

We hypothesized that the stories labeled original accounts (highest authenticity) would trigger more cognitive, affective, and physical involvement with the learning material (H1) and would foster higher situational interest and sense of insight (H2a) than the accounts labeled eyewitness accounts played by actors (medium authenticity) or reconstructed accounts (lowest authenticity). By contrast, we expected that participants who received the least authentic accounts would be more critical than participants who encountered the most authentic accounts (H2b). We expected that the different facets of authenticity would mediate the treatment effects on the processing and outcome variables (H3).

## 2. Method

The study was approved by the Ethics committee of the Faculty of Economics and Social Sciences at the University of Tübingen (A2.5.4-112\_aa) and preregistered on [aspredicted.org](https://aspredicted.org) ([https://aspredicted.org/blind.php?x=XMK\\_2ZJ](https://aspredicted.org/blind.php?x=XMK_2ZJ)).

### 2.1 Experimental Stimulus

We designed three experimental conditions that differed in authenticity framing by varying only the video's introduction, which was presented via text and audio simultaneously. All conditions included the information that the following video clip was produced for a documentary on the topic “children of war” (WWII) or “children of change” (GDR), respectively. We manipulated stimulus authenticity with regard to the person and the story. In the most authentic condition (*original*), we told participants they would watch a clip in which an eyewitness of the past would talk about past experiences (authentic person and authentic story). In the second condition with medium authenticity (*reenacted*), participants read and heard that the video would show an actor presenting someone else's actual past experience

(nonauthentic person, authentic story). In the third condition with low stimulus authenticity (*reconstructed*), we told participants that an actor would talk about reconstructed experiences from the past that, according to historians, could have happened in this way (nonauthentic person, nonauthentic content). In actuality, we did not change the videos at all—all clips showed “real” eyewitnesses—only the framing varied (see the Appendix for the exact wording of the framing).

We used a total of six clips. Three eyewitnesses (one woman, two men) talked about their experiences as children during WWII and after the war. The other three eyewitnesses (two women, one man), born in 1975, talked about their perspectives on the fall of the Berlin Wall. The materials thus differed by person, topic, and eyewitnesses' gender, which allowed us to statistically control for the effects of these features of the videos (see the Analysis section). The clips lasted from 7 to 10 min each and showed the person's head and upper body with a black background.

## 2.2 Participants and Design

We recruited participants via the online platform *prolific.org* and prescreened the sample for participants who either spoke German as their native language or were fluent in German and at the same time lived in a German-speaking country. Participants received 3.85 GBP for participating. Power analyses (*G\*Power* 3, Faul et al., 2007) revealed that a power of .80 can be achieved for the comparison of two equally sized groups assuming an effect size of small to medium size ( $d = .35$ ) at a sample size of at least 102 participants per group. We began with 452 participants; 24 dropped out before the randomization took place and the stimulus materials were shown, leaving a sample of  $N = 428$ . These 428 participants were randomly assigned by an algorithm implemented in *SoSci Survey* to the three treatment groups: 145 participants were in Condition 1 (original), 141 in Condition 2 (reenacted), and 142 in Condition 3 (reconstructed).

For the final sample, we checked whether participants understood the manipulation by asking two questions: Right after the clips, participants had to indicate whether they had seen an eyewitness or an actor. If they ticked actor, we asked if the story was reconstructed by historians or a real eyewitness' story. A second question, placed at the very end of the study, asked participants what introduction they have seen at the beginning of the video and showed the three possible screens as pictures. In total, 19.4% of the 428 participants did not pass one or both manipulation checks and were therefore excluded from further analyses (Original: 15.2%; Reenacted: 18.9%; Reconstructed: 23.2%). Furthermore, we excluded 10 participants who

stated that they had experienced severe technical problems. The final sample thus consisted of  $N = 336$  participants,  $n_{\text{original}} = 120$ ,  $n_{\text{reenacted}} = 110$ , and  $n_{\text{reconstructed}} = 106$ ; still sufficient to achieve a power of .80 as indicated by our power analysis. The mean age of the sample was  $M = 26.8$ ,  $SD = 5.7$  ( $M_{\text{original}} = 26.8$ ,  $SD = 5.5$ ;  $M_{\text{reenacted}} = 26.9$ ,  $SD = 6.0$ ;  $M_{\text{reconstructed}} = 26.8$ ,  $SD = 5.6$ ). In total, 58.3% of the participants stated they were female (Original: 56.7%; Reenacted= 59.1%; Reconstructed = 59.4%), 39.9% male (Original: 41.7%; Reenacted: 40.0%; Reconstructed: 37.7%), and 183 participants were enrolled at a university (Original: 65; Reenacted: 63; Reconstructed: 55).

### 2.3 Measures

**Situational processing and involvement.** Situational processing and involvement was assessed with a validated questionnaire (Author(s), in press) mapping the processes described by Zachrich et al. (2020). Table 1 presents scale statistics and sample items for all constructs.

**Table 1**

*Scale Descriptives for the Situational Processing and Involvement Constructs*

Construct	Number of items	$M$	$SD$	Cronbach's $\alpha$	Example item (translated)
<i>Cognitive processing and involvement</i>					
Attentional Focus	3	4.97	0.93	.81	I could fully concentrate on the story.
Vivid Imagination	3	4.57	1.15	.90	I could vividly picture the story in my head.
Abstract Imagination	3	4.44	1.02	.85	I could imagine the time the person talked about.
Perspective Recognition	4	4.72	0.93	.88	I could understand the person's thoughts.
Contextualization	3	4.77	1.07	.87	I linked the information to the knowledge that I already had.
<i>Affective processing and involvement</i>					
Being Moved	3	3.65	1.32	.92	The story moved me.
Unity of the Moment	3	3.27	1.21	.88	Encountering the story was special to me.
Immediacy	3	3.97	1.19	.86	The past suddenly felt very close.
Empathy	2	4.17	1.34	.88	I felt empathy for the person in the video.
Similarities	3	2.14	1.13	.84	Some of the experiences seemed similar to my own experiences.

*Physical processing and involvement*

Sensual Perception	2	3.41	1.30	.71	Although I only saw a video clip, it seemed that I could feel the person.
Physiological Response	2	1.76	1.07	.70	I physically reacted to the story.

*Note.* Item and scale analyses were conducted with the final sample of  $N = 336$ . All items were rated on a 6-point response format (1 = *completely disagree* to 6 = *completely agree*).

**Perceived authenticity.** Following Hofer (2016) and Hampp and Schwan (2014), we assessed six different facets of perceived authenticity as potential mediators, namely, Trustworthiness, Narrative Consistency, Plausibility, Typicality, Vividness of History, and Manipulation. Due to high correlations among some of the constructs (e.g., Trustworthiness with Plausibility:  $r = .82$ ) and weak model fit (RMSEA = .098; CFI = .876; SRMR = .135),  $\chi^2(df) = 449.91(64)$ ,  $p < .001$ , we conducted an exploratory factor analysis. Results suggested that two items be dropped due to double loadings, that Trustworthiness and Plausibility be combined into one factor, and that Narrative Consistency be included in the Vividness of History factor. Confirmatory factor analysis revealed a good fit of the model (RMSEA = .046; CFI = .978; SRMR = .052),  $\chi^2(df) = 101.38(59)$ ,  $p < .001$ . Table 2 presents the final scale statistics and example items. The Trust facet indicates whether participants rated the video clip as trustworthy and authentic. The Typicality facet indicates how typical the experiences in the clip seemed to many people. The Vividness of History facet indicates whether participants perceived that the personal story made the past comprehensible and come alive. Finally, the Manipulation facet indicates participants' ratings of interviewers' and producers' influence on the story and the recipient. As opposed to the first three facets, high ratings on this facet indicate that the story was not perceived as authentic.



**Table 2***Scale Descriptives for the Perceived Authenticity Facets*

Construct	Number of items	<i>M</i>	<i>SD</i>	Cronbach's $\alpha$	Example item (translated)
Trust	2	5.05	1.03	.87	The story was trustworthy to me.
Typicality	3	4.10	1.10	.88	The experiences in the video clip are typical for many people who lived during that time.
Vividness of History	5	4.68	0.88	.89	The personal story made the past come alive.
Manipulation	3	3.37	1.06	.70	The producer of the documentary wanted me to feel something specific when I watch the video clip.

*Note.* Item and scale analyses were conducted with the final sample of  $N = 336$ . All items were rated on a 6-point response format (1 = *completely disagree* to 6 = *completely agree*).

**Learning outcomes.** Table 3 presents the scale descriptives for the learning outcomes. We assessed participants' situational interest in history, their sense of insight into past lives (Author et al., in press), and their critical distance from personal stories of the past in general via an instrument assessing insights into the subjectivity of eyewitness accounts (Bertram et al., 2014). For Critical Distance, we used a partial credit model (Masters, 1982 scoring: 0, -1, -2) and weighted likelihood estimates (WLEs; Warm, 1989) as person parameter estimates, which showed an acceptable person separation reliability. We asked participants what they remembered of the story in a free text field. We were interested in whether they stored the information differently in the different treatment conditions. We analyzed the answers with an automatic dictionary approach using the Linguistic Inquiry and Word Count (LIWC; Meier et al., 2018; Pennebaker et al., 2015) software. We focused on three variable sets: (1) words marking cognitive processes, particularly (1a) words marking certainty (i.e., always;  $M = 2.74$ ,  $SD = 2.40$ ) and (1b) words marking tentativeness (i.e., maybe;  $M = 2.30$ ,  $SD = 2.23$ ); (2) words marking (2a) affective (i.e., cry;  $M = 3.39$ ,  $SD = 3.01$ ) and (2b) perceptual processes (i.e., hear;  $M = 1.80$ ,  $SD = 2.04$ ); and (3) linguistic dimensions, particularly (3a) total word count ( $M = 82.93$ ,  $SD = 61.65$ ) and (3b) question marks ( $M = 0.06$ ,  $SD = 0.41$ ). LIWC counts the relative number of words in a category related to the text.

**Table 3**  
*Scale Descriptives for the Learning Outcome Constructs*

Construct	Number of items	<i>M</i>	<i>SD</i>	Reliability	Response format	Example item (translated)
Situational Interest	6	2.84	0.76	$\alpha = .96$	4-point	I would like to know more about the topic from the account.
Sense of Insight	5	4.06	0.97	$\alpha = .90$	6-point	Through the story, I understood how it was back then.
Critical Distance	10	0.83	1.20	WLE-PSR = .76	partial credit, 0-2 points	Statements of eyewitnesses of the past should be compared with other sources and accounts.

*Note.* Item and scale analyses were conducted with the final sample of  $N = 336$ .  $\alpha =$  Cronbach's alpha; WLE-PSR = weighted likelihood estimate person separation reliability.

**Covariates.** We additionally assessed self-reported prior knowledge on the topic in focus with one item for each topic on a 4-point scale (1 = *knowing very little* to 4 = *knowing a lot*;  $M_{\text{GDR}} = 2.38$ ,  $SD_{\text{GDR}} = 0.62$ ;  $M_{\text{WWII}} = 2.80$ ,  $SD_{\text{WWII}} = 0.61$ ) and participants' dispositional interest in history with 11 items rated on a 4-point response format ( $M = 2.90$ ,  $SD = 0.66$ , Cronbach's  $\alpha = .94$ ).

## 2.5 Experimental Procedure

The study was conducted as an online experiment. After answering questions about prior knowledge and dispositional interest in history, participants read the introduction to the stimulus material. They were instructed to watch the video carefully and told that afterward they would be asked about their perception of the story and the information they remembered. Then participants were randomly assigned to one of three conditions. The framing was shown and read aloud within the first seconds of the video. After the clip was shown, participants answered the manipulation check and were then asked to rate the items on processing and involvement in the account. They were further asked to explain what they remembered from the clip in a free text field before rating their situational interest, perceived authenticity, and Critical Distance. The experiment ended with questions about personal data, technical problems and a final manipulation check item. Last, we debriefed participants and allowed them to withdraw their data, but no one did. The study was programmed in SoSci Survey. Participants took around 40 min on average to complete the study.

## 2.6 Data Analysis

We examined measurement models of the latent variables representing processing of the account and checked for measurement invariance across the groups as a first step. We used  $\Delta\text{CFI}$  to compare the measurement models and accepted the more restrictive model when  $\Delta\text{CFI} < .01$  (Cheung & Rensvold, 2002).  $\Delta\text{CFI}$  values were below .01 and thus indicated that strong measurement invariance could be assumed for every latent variable (see the Supplementary Online Material).

For RQ1 and RQ2, we then specified multiple-group models with one group for each of the three conditions. We included covariates (gender, age, final high-school GPA, dispositional interest in history) to improve power and to control for random group differences on the covariates. The covariates were centered on the grand mean (across all conditions), and the respective outcome was regressed on the covariates in each group with freely estimated regression weights in each group to account for potential Treatment  $\times$  Covariate interactions. Here, the intercepts represented expected values for participants with average values on the covariates and can be interpreted as adjusted means. We used likelihood ratio tests for models with freely estimated intercepts (Model 0) with models with equality constraints on group-specific intercepts (Model 1: original and reenacted; Model 2: original and reconstructed). We estimated adjusted mean differences between the groups through the Model Constraint option in Mplus 8 (Muthén & Muthén, 1998-2021).

Because the experimental material contained six different eyewitness accounts and two historical topics, we additionally checked for the possibility that the specific stimulus material—the two historical topics or the six different video clips—played a significant role in the treatment effects.

For RQ3, we specified mediation models with four facets of perceived authenticity as mediators of the effect of treatment on processing and on the outcome variables separately for the dependent variables that showed statistically significant group differences as investigated in RQ1 and RQ2 (MacKinnon, 2008). We included the treatment variable as a dummy variable (Model 1: 0 = original, 1 = reenacted; Model 2: 0 = original, 1 = reconstructed) and the four mediators. We used manifest scale scores instead of latent variables for the mediators to achieve higher statistical power in the structural part of the model, which was our focus here (Rhemtulla, 2016). The models also included the covariates as in the prior models. We used bootstrapping (10,000 bootstrap samples; MacKinnon, 2008) to ensure accurate statistical inference on the indirect effects. Additionally, we compared  $R^2$  in the model with all mediators with a model

without mediators to gauge how much the perceived authenticity facets contributed to the explained variance.

Analyses were computed in Mplus 8 (Muthén & Muthén, 1998-2021). One-sided significance tests were conducted on the .05 level according to our directed hypotheses (H1, H2a, H2b) in RQ1 and RQ2 by interpreting the sign of the parameter estimate and halved  $p$ -values. Scaling of the Critical Distance measure was done in ConQuest (Wu et al., 2007). For continuous predictor variables, we report standardized results that are based on both standardized predictor and outcome variables (STDYX section of the Mplus output). For binary predictors (e.g., the dummy-coded treatment variable) and their potential indirect effects, we report standardized results based on standardized outcomes only (STDY section of the Mplus output).

Our statistical analyses were in line with the preregistered approach (see [https://aspredicted.org/blind.php?x=XMK\\_2ZJ](https://aspredicted.org/blind.php?x=XMK_2ZJ)), but some extensions should be noted. First, we decided to use an automatic assessment of the outcome remembering the information, something that had not been preregistered. Second, we conducted additional analyses to strengthen the study by exploring the measurement invariance of the latent variables across groups and by running a number of robustness checks. Third, for RQ3, the analytic procedure had not been preregistered in detail.

### 3. Results

#### 3.1 RQ1: The More Authentic, the More (Emotionally) Involved?

Table 4 presents the results of the analyses of treatment condition differences for all latent processing and involvement variables. Likelihood difference tests of models with equality constraints on group intercepts and models with freely estimated intercepts indicated that the latter fit statistically significantly better than those with equality constraints across group intercepts for 2 out of 10 group comparisons for cognitive involvement, for 7 out of 10 group comparisons for affective involvement, and for 2 out of 4 group comparisons for physical involvement.

The direction of the mean differences confirmed our assumption that participants were more deeply involved with the stories that were framed as more authentic than those that were framed as less authentic (except for Similarities in both group comparisons), although not all of the mean differences were statistically significant. Results revealed statistically significant differences in two cognitive processing constructs, Vivid Imagination and Contextualization, when comparing the original and reconstructed conditions. For the affective processing and

involvement constructs, we found consistent statistically significant differences for the original versus reconstructed conditions (except Similarities). Additionally, Being Moved, Unity of the Moment, and Empathy also showed statistically significant differences between the original and reenacted conditions. Furthermore, results revealed a statistically significant difference for two physical processing constructs: Physiological Response between the original and reenacted conditions and Sensual Perception between the original and reconstructed conditions.

In summary, watching videos labeled “original” versus “reenacted” or “reconstructed by historians” especially affected affective processing and involvement with the story and some cognitive and physical processes. Specifically, the more authentic the story was described to be, the more emotionally involved the participants were.

### 3.2 RQ 2: Authenticity Framing and Learning Outcomes

With regard to the outcome measures (see Table 4), likelihood difference tests indicated that there would be statistically significant group differences for one of the three outcome measures. In fact, participants’ situational interest in the original condition was statistically significantly higher than in the reenacted and reconstructed conditions. However, contrary to our a priori expectations, analyses revealed no differences for participants’ feeling of knowing what the past must have been like in the original condition in comparison with participants in the reenacted and reconstructed conditions. The adjusted mean Critical Distance differences were in the expected direction (higher in the reconstructed and reenacted than in the original) but were not statistically significant.

Comparing participants’ word use in the free text field that asked what participants remembered from the story, we did not find statistically significant differences (two-sided  $p$ -values) for any of the variable sets we investigated or for either group comparison. Adjusted mean differences ranged from  $-0.44 < \Delta M < 0.62$ , all  $ps > .110$ , for the variable sets affective and cognitive processes. The total word count was also not statistically significantly different between conditions (1 vs. 2:  $\Delta M = 6.18$ ,  $SE = 8.37$ ,  $p = .460$ ; 1 vs. 3:  $\Delta M = -4.75$ ,  $SE = 7.29$ ,  $p = .515$ ).

In summary, in line with our expectations, we found effects of the authenticity framing on situational interest. Conversely, we found no statistically significant differences across experimental conditions for Sense of Insight or Critical Distance. Also, remembering the information did not seem to differ by authenticity framing condition.

**Table 4**

*Group Differences in Adjusted Means of the Latent Constructs Measuring Situational Processing and Involvement and Learning Outcomes*

	Original versus reenacted			Original versus reconstructed		
	$\Delta M$	<i>SE</i>	<i>p</i>	$\Delta M$	<i>SE</i>	<i>p</i>
<i>Cognitive processing and involvement</i>						
Attentional Focus	-0.12	0.12	.178	-0.23	0.13	.052
Vivid Imagination	-0.09	0.10	.156	<b>-0.29</b>	<b>0.12</b>	<b>.002</b>
Abstract Imagination	-0.06	0.13	.322	-0.18	0.14	.110
Perspective Recognition	0.01	0.10	.453	-0.18	0.12	.050
Contextualization	-0.12	0.12	.167	<b>-0.25</b>	<b>0.14</b>	<b>.025</b>
<i>Affective processing and involvement</i>						
Being Moved	<b>-0.33</b>	<b>0.16</b>	<b>.021</b>	<b>-0.52</b>	<b>0.17</b>	<b>.003</b>
Unity of the Moment	<b>-0.55</b>	<b>0.15</b>	<b>&lt;.001</b>	<b>-0.79</b>	<b>0.16</b>	<b>&lt;.001</b>
Immediacy	-0.18	0.11	.056	<b>-0.28</b>	<b>0.13</b>	<b>.025</b>
Empathy	<b>-0.27</b>	<b>0.15</b>	<b>.040</b>	<b>-0.38</b>	<b>0.16</b>	<b>.005</b>
Similarities	0.04	0.14	.381	0.08	0.16	.300
<i>Physical processing and involvement</i>						
Sensual Perception	-0.14	0.09	.071	<b>-0.35</b>	<b>0.17</b>	<b>&lt;.001</b>
Physiological Response	<b>-0.26</b>	<b>0.16</b>	<b>.032</b>	-0.22	0.16	.097
<i>Learning outcomes</i>						
Sense of Insight	-0.15	0.10	.081	-0.14	0.12	.113
Situational Interest	<b>-0.15</b>	<b>0.08</b>	<b>.034</b>	<b>-0.16</b>	<b>0.08</b>	<b>.026</b>
Critical Distance	0.05	0.14	.368	0.10	0.16	.267

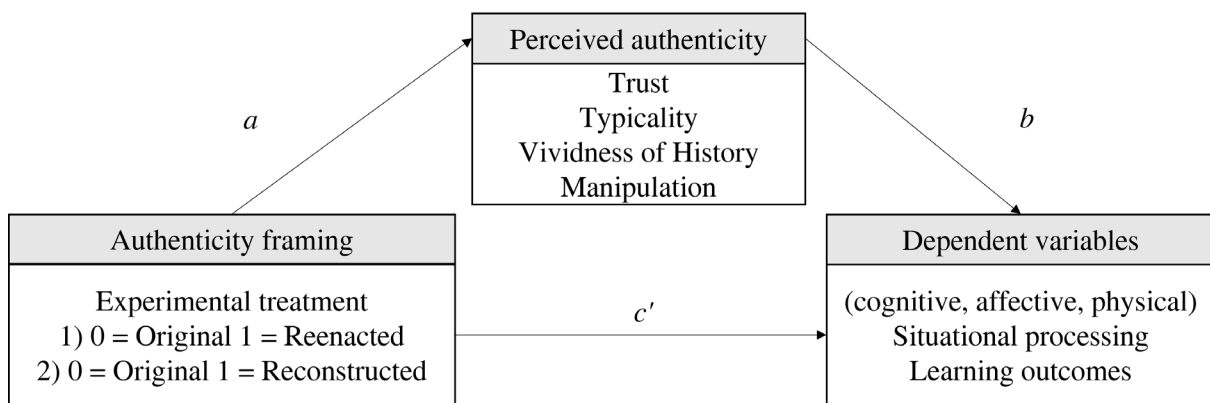
*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reenacted}} = 110$ ;  $n_{\text{reconstructed}} = 106$ . The original group served as the reference group. Adjusted means refer to the expected values for participants with average values on the covariates.  $\Delta M$  = Adjusted mean difference (Reenacted - Original, Reconstructed - Original), *SE* = standard error, *p* = *p*-values of likelihood difference tests for models with freely estimated intercepts across groups versus models with equality constraints on the intercepts. One-tailed *p*-values are reported. Statistically significant coefficients are printed in bold.

### 3.3 RQ3: The (Mediating) Role of Perceived Authenticity

Our mediation model for RQ3 is graphically depicted in Figure 2. Mediation models are based on the assumption that the effect of a predictor on a dependent variable can be explained via one or more mediators (paths  $a$  and  $b$  for a single mediator model). Thus, an indirect effect of the predictor on the dependent variable via the mediator variable(s) (paths  $ab$ ) is assumed, whereas the direct effect(s) (path  $c'$ ) should ideally be zero, indicating complete mediation. We investigated mediation effects for the dependent variables that showed statistically significant group differences as reported in 3.1 and 3.2.

**Figure 2**

*Mediation Model*



Looking at path  $a$  in our models, we found that only two out of eight effects were statistically significant. More specifically, participants rated the story as more trustworthy in the original than in the reconstructed condition ( $a_1 = -0.48$ ,  $SE = 0.13$ ,  $p < .001$ ). Similarly, participants in the original condition rated the account as less manipulated by the producer of the clip than participants in the reconstructed condition ( $a_4 = 0.47$ ,  $SE = 0.13$ ,  $p < .001$ ).

For path  $b$ , 21 of the 52 paths we investigated were statistically significant. Of interest, one facet, Vividness of History, was a consistent statistically significant predictor of all the dependent variables. Additionally, we found a substantial increase in the variance explained in the dependent variables when we compared models with all four facets of perceived authenticity as mediators with regression models without mediators. The increase in explained variance ( $\Delta R^2$ ) ranged from .13 to .52, which highlights the potentially important role of perceived authenticity.

However, out of the 52 indirect effects (paths  $ab$ ) we analyzed, only one (effect of the original vs. the reconstructed condition on Unity of the Moment via Trust;  $ab = 0.095$ ,  $SE = 0.41$ ,  $p = .022$ ) was statistically significant; the direct effect (path  $c'$ ) remained statistically

significant. In sum, we did not find support for our mediation hypotheses. However, the substantial increase in explained variance when adding perceived authenticity to the models documents the predictive power of perceived authenticity. The complete tables containing paths  $b$ ,  $c'$ , and  $ab$  are included in the Supplementary Online Material.

### **3.4 Robustness Checks: The Roles of the Topic and Story Tellers' Unique Experiences**

As robustness checks, we investigated how much the results reported for RQ1 or RQ2 were affected by differences across the six eyewitness accounts or the two historical topics covered by the videos. Results are reported in detail in the Supplementary Online Material. In short, although there were some modest differences in the effect sizes across the six clips/two historical topics, the effects all had the same and expected direction and generally concurred with the findings reported here. In sum, these additional analyses attest to the robustness of the results across experiences and topics and, hence, the generalizability of the findings.

## **4. Discussion**

The current study was aimed at opening the black box of the mechanisms underlying authentic learning by investigating the effects of authentic learning material on situational processing and involvement as well as the mediating role of perceived authenticity in this context. The authenticity of the learning material—videos about personal stories of the past—was varied by solely varying the framing of the video clip as (a) original, (b) reenacted by actors, or (c) reconstructed by historians and played by actors. There were three main results. First, in line with our hypotheses, the three experimental conditions led to substantially different situational processing and involvement; the strongest and most consistent effects occurred for affective involvement, where four out of five facets differed across experimental conditions. Second, again in line with our hypothesis, we also found a positive effect of the most authentic report (original condition) on situational interest; however, the framing did not affect the other constructs under investigation (feeling of knowing what the past must have been like, critical understanding of personal accounts in general, linguistic features of remembering). Third, although we did not find support for our hypothesis that effects of the differential framing of the reports about the past (the three experimental conditions) on situational processing and the outcome variables would be mediated by perceived authenticity, there was nevertheless a clear indication that perceived authenticity matters a lot. The amount of variance explained in the dependent variables substantially increased when the mediators were included, and particularly the Vividness of History facet consistently showed predictive power for all the situational processing and involvement variables. We next discuss how the results contribute to a better



understanding of learning with personal stories of the past in history education and their implications for authentic learning more generally.

#### **4.1 Authentic Learning in History Education**

Many people associate learning about history and the past with boring activities, such as learning dates or reading chronicles. Authentic learning offers the opportunity to understand what history has to do with the present and one's daily life—a connection that learners do not necessarily understand at first sight (Straaten et al., 2016). Eyewitness accounts—personal stories of the past—are authentic learning materials that bridge the past and present and are therefore motivating learning materials (Bertram et al., 2017). Moreover, our study hints that not only might the authentic stories be motivating and foster interest, but they can also trigger affective processing and involvement with the material. Learners are touched by the story and feel empathy or an immediate connection to the past—not only when learners believed that the story and the person were original, but, even more important, when they perceived the personal story was suitable for providing valuable insights into the past, making it come alive and comprehensible.

Whereas researchers in the field of history teaching now agree that emotional processes play an important role in the learning process, the interconnectedness of emotions in the learning context is not yet clear (e.g., Brauer & Lücke, 2021). The current study makes two important contributions. First, it increases the understanding of what makes up the authenticity of eyewitness accounts. Framing the accounts as original versus reconstructed affected perceived trustworthiness. Moreover, we were able to differentiate between several facets of perceived authenticity. Second, our study gives insight into how authentic stories of the past affect different learning and involvement processes: the authenticity framing and between-person differences in perceived authenticity affected learners' cognitive, affective, and physical processing and involvement with the story.

#### **4.2 Extending the Conceptualization of Authentic Learning**

Whereas a general belief has evolved that authentic learning settings boost students' learning outcomes, empirical research has drawn an inconsistent picture of the effects of authentic learning. Our study demonstrates that the effects of authentic learning might be more complex than originally envisioned. More specifically, our empirical results have two main implications. First, in line with a similar postulate by Schwan et al. (2017) and as Figure 1 shows, theoretical accounts of what happens in authentic learning should explicitly take into account the role of situational processing and involvement in the learning process. Second, our

study also suggests the need to get a better understanding of the role of different conceptualizations of authenticity. As our results showed, both the authenticity framing and participants' perception of authenticity (as measured by four different facts) mattered for how they perceived the videos.

For involvement processes triggered by authentic learning materials, affective processing and involvement stood out. This pattern of findings is generally in line with prior research that found positive effects of authentic learning on motivation (Bertram et al., 2017; Itzek-Greulich et al., 2017). At the same time, given that affective processing can prevent a learner from transferring their knowledge to the learning goal, one may ask to what extent affective involvement processes may have contributed to the lack of effects of authentic learning on achievement outcomes observed in some prior studies (e.g., Itzek-Greulich et al., 2015; Riegel & Kindermann, 2016).

In our study, involvement processes were affected not only by the actual authenticity framing but by different facets of perceived authenticity. These findings extend the conceptualization by Fougat et al. (2019), who differentiated between different dimensions of actual authenticity by stressing that not every learner seems to be triggered by the actual authenticity framing. Even more important is the perception of the learning material as more or less authentic on different facets. Hence, our study supports the conceptualization of perceived authenticity as the result of Betz et al.'s (2016) attributional process, which evolves from the interplay between the learning material (and its authenticity) and learners' characteristics.

The Vividness of History facet, which consistently predicted situational processing and involvement variables as well as situational interest, captured participants' ratings of whether the personal story was suitable for making history come alive and comprehensible. More generally speaking, this facet indicated whether the authentic learning material contributed to a better understanding of the learning content and to achieving the learning goal. Hence, how learners rate the authentic learning materials or setting with regard to its utility in the learning process seem to be especially important.

### **4.3 Limitations and Future Research**

Although this study contributes to a more sophisticated understanding of authentic learning, it comes with limitations. First, we used a total of six video clips, and thus, it is possible that different results would have been found with other clips. Of course, future research is needed to address this issue. However, in an extensive set of robustness checks, the general pattern of results remained remarkably similar across the six video clips.

---

Second, our sample consisted of adult participants recruited via a commercial online platform. This sample composition might trigger questions about the generalizability of our findings. A first question pertains to the participants' ages: Although authentic learning is believed to positively affect learning across a wide age range (e.g., Merkt et al., 2020), more research is necessary in this regard. Second, it is unclear to what extent the online format of the study affected participants' attention levels or whether a more formal setting (e.g., a classroom) would make any difference. Still, our manipulation check showed that the large majority of participants paid close attention to the information we provided.

Finally, the length of the study (experimental stimulus varying from 7 to 10 min), its online character, and a lack of instructional embedding prevented us from investigating additional achievement outcomes that would attest to the effectiveness of authentic learning. However, at this point, the strength of this experimental study can be seen in its extension of the understanding of authentic learning by looking at situational processing and involvement variables and differentiating between different facets of perceived authenticity. Future research might include randomized intervention studies in real classrooms that also allow for interactions among participants and a wider array of topics.

#### **4.4 Conclusion**

This experimental online study adds to the research field of authentic learning by examining the effects of authentic learning material on situational processing and involvement and not only outcome variables. Our study indicates that especially affective involvement is affected by the authentic learning material and that authenticity has more facets than only its real-world character. Our study suggests directions for further research and has an implication for the design of authentic learning materials—which is in line with prior research: Authenticity is not a value in itself but needs to be considered within the interplay of content and learning goals.

---

### References

- Author(s) (in press). [details removed for peer review].
- Ballis, A., & Schwendemann, L. (2021). 'In any case, you believe him one hundred percent, everything he says.' Trustworthiness in Holocaust survivor talks with high school students in Germany. *Holocaust Studies*, 28(2), 191–220.  
<https://doi.org/10.1080/17504902.2021.1915016>
- Bertram, C., Wagner, W., & Trautwein, U. (2017). Learning historical thinking with oral history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54(3), 444–484.  
<https://doi.org/10.3102/0002831217694833>
- Betz, A. (2018). The influence of the learning environment on learners' (perceived) authenticity of science communication and on their situational interest. *Unterrichtswissenschaft*, 46(3), 261–278. <https://doi.org/10.1007/s42010-018-0021-0>
- Betz, A., Flake, S., Mierwald, M., & Vanderbeke, M. (2016). Modelling authenticity in teaching and learning contexts: A contribution to theory development and empirical investigation of the construct. *Proceedings of International Conference of the Learning Sciences, ICLS*, 2, 815–818.
- Bilandzic, H., & Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, 36(1), 29–50.  
<https://doi.org/10.1515/COMM.2011.002>
- Brauer, J., & Lücke, M. (2021). Emotionen, Geschichte und historisches Lernen. In Gläser-Zirkuda, Hofmann, & V. Frederking (Eds.), *Emotionen im Unterricht*. Kohlhammer.
- Brüning, C. I. (2018). *Holocaust Education in der heterogenen Gesellschaft*. Wochenschau.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.  
[https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Cronin, J. F. (1993). Four misconceptions about authentic learning. *Educational Leadership*, April, 78–80.
- Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology*. Springer.
- De Jong, S. (2018). *The witness as object--video testimony in memorial museums*. Berghahn Books.
- Feddermann, M., Möller, J., & Baumert, J. (2021). Effects of CLIL on second language learning: Disentangling selection, preparation, and CLIL-effects. *Learning and Instruction*, 74. <https://doi.org/10.1016/j.learninstruc.2021.101459>

- Fougts, S. S., Misfeldt, M., & Shaffer, D. W. (2019). Realistic authenticity. *Journal of Interactive Learning Research*, 30(4), 477–504.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721.  
<https://doi.org/10.1037/0022-3514.79.5.701>
- Hampp, C., & Schwan, S. (2014). Perception and evaluation of authentic objects: Findings from a visitor study. *Museum Management and Curatorship*, 29(4), 349–367.  
<https://doi.org/10.1080/09647775.2014.938416>
- Henke-Bockschatz, G. (2014). *Oral History im Geschichtsunterricht [Oral history in history lessons]*. Wochenschau.
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23–48.  
<https://doi.org/10.1007/BF02319856>
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2015). Effects of a science center outreach lab on school students' achievement. *Learning and Instruction*, 38, 43–52. <https://doi.org/10.1016/j.learninstruc.2015.03.003>
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2017). Effectiveness of lab-work learning environments in and out of school: A cluster randomized study. *Contemporary Educational Psychology*, 48, 98–115.  
<https://doi.org/10.1016/j.cedpsych.2016.09.005>
- Latham, K. F. (2013). Numinous experiences with museum objects. *Visitor Studies*, 16(1), 3–20. <https://doi.org/10.1080/10645578.2013.767728>
- Lombardi, M. M. (2007). *Authentic learning for the 21st century: An overview* (D. G. Oblinger (ed.); Vol. 1, pp. 1–7). EDUCAUSE Learning Initiative.  
<http://alicechristie.org/classes/530/EduCause.pdf>
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Erlbaum.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–147. <https://doi.org/https://doi.org/10.1007/BF02296272>
- Meier, T., Boyd, R. L., Pennebaker, J. W., Mehl, M. R., Martin, M., Wolf, M., & Horn, A. B. (2018). “LIWC auf Deutsch”: The development, psychometrics, and introduction of DE-LIWC2015. <https://osf.io/tfqzc/>
- Merkt, M., Lux, S., Hoogerheide, V., Gog, T. Van, & Schwan, S. (2020). A change of scenery: Does the setting of an instructional video affect learning? *Journal of Educational Psychology*, 112(6), 1273–1283.

- <https://doi.org/https://doi.org/10.1037/edu0000414>
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Nachtigall, V., & Rummel, N. (2021). Investigating students' perceived authenticity of learning activities in an out-of-school lab for social sciences: a replication study. *Instructional Science*, *49*(6), 779–810. <https://doi.org/10.1007/s11251-021-09556-3>
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015* (pp. 1–26). The University of Texas at Austin.
- Rhemtulla, M. (2016). Population performance of SEM parceling strategies under measurement and structural model misspecification. *Psychological Methods*, *21*(3), 348–368. <https://doi.org/https://doi.org/10.1037/met0000072>
- Riegel, U., & Kindermann, K. (2016). Why leave the classroom? How field trips to the church affect cognitive learning outcomes. *Learning and Instruction*, *41*, 106–114. <https://doi.org/10.1016/j.learninstruc.2015.10.004>
- Rule, A. C. (2006). Editorial: The components of authentic learning. *Journal of Authentic Learning*, *3*(1), 1–10.
- Sabrow, M. (2012). Der Zeitzeuge als Wanderer zwischen den Welten. In M. Sabrow & N. Frei (Eds.), *Die Geburt des Zeitzeugen nach 1945* (pp. 13–32). Wallstein.
- Schwan, S., Bauer, D., Kampschulte, L., & Hampp, C. (2017). Representation equals presentation? Photographs of objects receive less attention and are less well remembered than real objects. *Journal of Media Psychology*, *29*(4), 176–187. <https://doi.org/10.1027/1864-1105/a000166>
- Shaffer, D. W., & Resnick, M. (1999). “Thick” authenticity: New media and authentic learning. *Journal of Interactive Learning Research*, *10*(2), 195–215. <https://doi.org/10.1016/j.stomax.2007.01.005>
- Stamer, I., David, M. A., Höffler, T., Schwarzer, S., & Parchmann, I. (2021). Authentic insights into science: Scientific videos used in out-of-school learning environments. *International Journal of Science Education*, *43*(6), 868–887. <https://doi.org/10.1080/09500693.2021.1891321>
- Straaten, D. Van, Wilschut, A., & Oostdam, R. (2016). Making history relevant to students by connecting past, present and future: a framework for research. *Journal of Curriculum Studies*, *0272*, 1–24. <https://doi.org/10.1080/00220272.2015.1089938>
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427–450. <https://doi.org/doi:10.1007/BF02294627>

- 
- Whitman, G. (2004). *Dialogue with the past*. Altamira.
- Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER ConQuest version 2.0: Generalized item response modelling software*. ACER.
- Yıldız, E. P. (2021). Analysis of the last 10 years of articles and theses on authentic learning: A meta analysis study. *African Educational Research Journal*, 9(2), 505–514.  
<https://doi.org/10.30918/aerj.92.21.068>
- Zachrich, L., Weller, A., Baron, C., & Bertram, C. (2020). Historical experiences: A framework for encountering complex historical sources. *History Education Research Journal*, 17(2), 243–275.

## **Appendix**

### **Experimental Framing**

#### **1. Introduction for all groups**

Different experiences of people who were still children at the time of World War II and teenagers at the time of the fall of the Berlin Wall were presented for two TV documentaries with the titles “Children of War” and “Children of Change.” The goal of the documentaries was to bring the past vividly to life through personal life stories.

Three types of personal accounts were used in the documentary:

- (1) Original report: The eyewitnesses told about their personal experiences in an authentic report.
- (2) Reenacted report: Some interviews were available only as audio recordings. These authentic accounts were acted out by actors as if the experiences were their own.
- (3) Reconstructed report: In order to be able to add further typical experiences, historians reconstructed life stories. These reconstructed accounts were also played by actors as if the experiences were their own.

In the following, you will see **only one** video clip of these three types of personal accounts.

#### **2. Framing within the video clip (using the example of "Children of War," male person)**

##### *Group original*

In the following video clip, an eyewitness describes his personal experiences in the past for a documentary on the topic of “Children of War.”

The eyewitness was asked to simply recount the experiences as he experienced them.

##### *Group reenacted*

In the following video clip, an actor describes the personal experiences of an eyewitness for a documentary on the topic of “Children of War,” of which only an audio recording was available.

The actor was asked to tell the story as if he had experienced it himself.

##### *Group reconstructed*

In the following video clip, an actor describes reconstructed experiences for a documentary on the topic of “Children of War,” which, according to historians, were experienced in this or a similar way by a person in the past.

The actor was asked to tell the reconstructed story as if he had experienced it himself.



---

**Supplementary Material****Table of Contents**

1. Measurement Invariance	S. 143
2. Mediation Analyses	S. 147
3. Robustness Checks	S. 151

**1. Measurement Invariance**

We examined the measurement models of the latent variables that referred to participants' processing of the account and checked for measurement invariance across the groups as a first step. For configural measurement invariance, the factor structure of the latent variable was set equal across groups; for weak measurement invariance, the factor loadings were additionally set equal across groups; for strong measurement invariance, we additionally set the intercepts of the groups equal. We used  $\Delta\text{CFI}$  to compare the measurement models and accepted the more restrictive model when  $\Delta\text{CFI} < .01$  (Cheung & Rensvold, 2002).  $\Delta\text{CFI}$  values were below .01 and thus indicated that strong measurement invariance could be assumed for every latent variable. Goodness-of-fit statistics are presented in Tables 1.1 to 1.3.

**Table 1.1**

*Goodness-of-Fit Statistics for Tests of Measurement Invariance for Latent Constructs of Cognitive Situational Processing and Involvement*

	RMSEA	CFI	SRMR	$\chi^2(df)$	<i>p</i>	$\Delta$ CFI
<i>Attentional focus</i>						
Configural invariance	0.027	0.995	0.031	32.37(30)	.351	
Weak invariance	0.027	0.994	0.043	36.82(34)	.340	0.001
Strong invariance	0.034	0.989	0.044	42.96(38)	.267	0.005
<i>Vivid imagination</i>						
Configural invariance	0.000	1.000	0.023	26.44(30)	.652	
Weak invariance	0.004	1.000	0.053	34.07(34)	.464	0.000
Strong invariance	0.000	1.000	0.058	35.31(38)	.595	0.000
<i>Abstract imagination</i>						
Configural invariance	0.050	0.983	0.028	38.49(30)	.138	
Weak invariance	0.043	0.986	0.037	41.06(34)	.189	-0.003
Strong invariance	0.035	0.989	0.037	43.22(38)	.258	-0.003
<i>Perspective recognition</i>						
Configural invariance	0.000	1.000	0.027	25.74(30)	.688	
Weak invariance	0.000	1.000	0.041	25.98(34)	.836	0.000
Strong invariance	0.000	1.000	0.041	26.46(38)	.921	0.000
<i>Contextualization</i>						
Configural invariance	0.000	1.000	0.025	27.70(30)	.586	
Weak invariance	0.000	1.000	0.030	26.72(34)	.809	0.000
Strong invariance	0.000	1.000	0.029	29.51(38)	.837	0.000

*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reenacted}} = 110$ ;  $n_{\text{reconstructed}} = 106$ . CFI values were used for the model comparisons. The more restrictive model could be concluded to hold when  $\Delta$ CFI < .01.

**Table 1.2**

*Goodness-of-Fit Statistics for Tests of Measurement Invariance for Latent Constructs of Affective Situational Processing and Involvement*

	RMSEA	CFI	SRMR	$\chi^2(df)$	<i>p</i>	$\Delta$ CFI
<i>Being moved</i>						
Configural invariance	0.009	1.000	0.018	30.30(30)	.450	
Weak invariance	0.023	0.998	0.028	36.00(34)	.375	0.002
Strong invariance	0.027	0.996	0.031	41.07(38)	.338	0.002
<i>Unity of the moment</i>						
Configural invariance	0.053	0.984	0.028	39.61(30)	.113	
Weak invariance	0.054	0.982	0.040	44.90(34)	.100	0.002
Strong invariance	0.054	0.980	0.039	50.34(38)	.087	0.002
<i>Immediacy</i>						
Configural invariance	0.040	0.991	0.031	35.42(30)	.228	
Weak invariance	0.033	0.993	0.040	38.26(34)	.282	-0.002
Strong invariance	0.038	0.990	0.043	44.04(38)	.993	0.003
<i>Empathy</i>						
Configural invariance	0.028	0.997	0.023	13.08(12)	.718	
Weak invariance	0.000	1.000	0.015	12.54(14)	.563	-0.003
Strong invariance	0.000	1.000	0.017	13.55(16)	.632	0.000
<i>Similarities</i>						
Configural invariance	0.081	0.954	0.035	51.99(30)	.008	
Weak invariance	0.073	0.957	0.041	54.40(34)	.015	-0.003
Strong invariance	0.064	0.963	0.040	55.57(38)	.033	-0.006

*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reenacted}} = 110$ ;  $n_{\text{reconstructed}} = 106$ . CFI values were used for the model comparisons. The more restrictive model could be concluded to hold when  $\Delta$ CFI < .01.

**Table 1.3**

*Goodness-of-Fit Statistics for Tests of Measurement Invariance for Latent Constructs of Physical Situational Processing and Involvement and Learning Outcomes*

	RMSEA	CFI	SRMR	$\chi^2(df)$	<i>p</i>	$\Delta$ CFI
<i>Physical processing and involvement</i>						
<i>Sensual perception</i>						
Configural invariance	0.000	1.000	0.015	6.16(12)	.908	
Weak invariance	0.000	1.000	0.024	13.12(14)	.514	0.000
Strong invariance	0.009	0.999	0.032	16.14(16)	.443	0.001
<i>Physiological response</i>						
Configural invariance	0.067	0.968	0.025	18.03(12)	.114	
Weak invariance	0.049	0.980	0.025	17.77(14)	.217	-0.012
Strong invariance	0.028	0.993	0.025	17.41(16)	.359	-0.013
<i>Learning outcomes</i>						
<i>Sense of insight</i>						
Configural invariance	0.000	1.000	0.022	19.77(30)	.922	
Weak invariance	0.000	1.000	0.027	21.15(34)	.958	0.000
Strong invariance	0.000	1.000	0.027	23.30(38)	.971	0.000
<i>Situational interest</i>						
Configural invariance	0.065	0.977	0.041	118.00(80)	.004	
Weak invariance	0.064	0.976	0.053	128.58(88)	.003	0.001
Strong invariance	0.059	0.978	0.054	132.93(96)	.008	-0.002

*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reenacted}} = 110$ ;  $n_{\text{reconstructed}} = 106$ . CFI values were used for the model comparisons. The more restrictive model could be concluded to hold when  $\Delta$ CFI < .01.

## 2. Mediation Analyses

Paths  $b$  and  $c'$  and  $ab$  as well as  $R^2$  and  $\Delta R^2$  of the respective mediation models are displayed in Tables 2.1 to 2.3.

**Table 2.1**

*Treatment Effects on Mediators (Facets of Perceived Authenticity) as well as Direct, Indirect, and Mediator Effects for Comparisons of the Original and Reenacted Conditions*

	Affective processing and involvement									Physical processing and involvement			Learning outcomes		
	Being moved			Unity of the moment			Empathy			Physiological response			Situational interest		
	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>
<i>Mediator effects on dependent variables</i>															
Trust ( $b_1$ )	<b>-0.18</b>	<b>0.08</b>	<b>.027</b>	-0.11	0.08	.161	-0.08	0.09	.359	-0.23	0.13	.071	0.03	0.09	.750
Typicality ( $b_2$ )	<b>0.14</b>	<b>0.06</b>	<b>.021</b>	0.08	0.06	.183	<b>0.16</b>	<b>0.06</b>	<b>.009</b>	0.12	0.08	.150	-0.11	0.07	0.12
History ( $b_3$ )	<b>0.67</b>	<b>0.09</b>	<b>&lt;.001</b>	<b>0.61</b>	<b>0.09</b>	<b>&lt;.001</b>	<b>0.62</b>	<b>0.10</b>	<b>&lt;.001</b>	<b>0.47</b>	<b>0.12</b>	<b>&lt;.001</b>	<b>0.50</b>	<b>0.09</b>	<b>&lt;.001</b>
Manipulation ( $b_4$ )	0.05	0.05	.338	<b>0.15</b>	<b>0.06</b>	<b>.012</b>	0.00	0.06	.991	<b>0.22</b>	<b>0.08</b>	<b>.004</b>	0.03	0.06	.557
<i>Direct effects</i>															
Treatment ( $c'$ )	<b>-0.26</b>	<b>0.11</b>	<b>.015</b>	<b>-0.42</b>	<b>0.11</b>	<b>&lt;.001</b>	-0.19	0.11	.075	<b>-0.34</b>	<b>0.15</b>	<b>.025</b>	-0.10	0.11	.354
<i>Indirect effects</i>															
$a_1b_1$	0.05	0.04	.196	0.03	0.03	.295	0.02	0.03	.450	0.05	0.15	.221	-0.01	0.03	.790
$a_2b_2$	0.02	0.03	.466	0.01	0.02	.555	0.02	0.02	.441	0.01	0.02	.544	-0.01	0.02	.514
$a_3b_3$	-0.02	0.10	.827	-0.02	0.08	.828	-0.02	0.08	.829	-0.01	0.06	.833	-0.01	0.06	.829
$a_4b_4$	0.01	0.02	.534	0.02	0.02	.309	0.00	0.12	.993	0.04	0.04	.309	0.01	0.01	.678
$R^2$	0.52	0.05		0.50	0.05		0.51	0.06		0.35	0.08		0.45	0.05	
$\Delta R^2$	0.35			0.34			0.37			0.20			0.21		

*Note.*  $n = 230$ ; *Est.* = Estimate of regression coefficient, *SE* = standard error, treatment is dummy-coded 0 = original, 1 = reenacted,  $\Delta R^2$  = difference compared with a model without mediators. Statistically significant coefficients at the  $p \leq .05$  level are printed in bold.

**Table 2.2**

*Treatment Effects on Mediators (Facets of Perceived Authenticity) as well as Direct, Indirect, and Mediator Effects for Comparisons of the Original and Reconstructed Conditions*

	Affective processing and involvement											
	Being Moved			Unity of the Moment			Immediacy			Empathy		
	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>
<i>Mediator effects on dependent variables</i>												
Trust ( $b_1$ )	<b>-0.16</b>	<b>0.08</b>	<b>.046</b>	<b>-0.20</b>	<b>0.07</b>	<b>.007</b>	0.00	0.08	.982	-0.08	0.09	.321
Typicality ( $b_2$ )	0.04	0.05	.428	0.03	0.06	.612	-0.04	0.06	.503	<b>0.12</b>	<b>0.06</b>	<b>0.05</b>
History ( $b_3$ )	<b>0.77</b>	<b>0.08</b>	<b>&lt;.001</b>	<b>0.71</b>	<b>0.08</b>	<b>&lt;.001</b>	<b>0.77</b>	<b>0.09</b>	<b>&lt;.001</b>	<b>0.69</b>	<b>0.09</b>	<b>&lt;.001</b>
Manipulation ( $b_4$ )	0.04	0.05	.421	0.08	0.06	.171	-0.04	0.05	.424	0.01	0.06	.902
<i>Direct effects</i>												
Treatment ( $c'$ )	<b>-0.34</b>	<b>0.11</b>	<b>.002</b>	<b>-0.57</b>	<b>0.11</b>	<b>&lt;.001</b>	-0.13	0.10	.195	-0.20	0.12	.091
<i>Indirect effects</i>												
$a_1b_1$	0.07	0.04	.073	0.10	0.04	.022	0.00	0.04	.983	0.04	0.04	.345
$a_2b_2$	0.00	0.01	.867	0.00	0.01	.902	0.00	0.01	.881	0.00	0.02	.809
$a_3b_3$	-0.16	0.10	.108	-0.15	0.09	.104	-0.16	0.10	.112	-0.14	0.09	.108
$a_4b_4$	0.02	0.03	.458	0.04	0.03	.211	-0.02	0.02	.443	0.00	0.03	.907
$R^2$	0.56	0.05		0.55	0.05		0.67	0.05		0.53	0.06	
$\Delta R^2$	0.42			0.34			0.52			0.42		

*Note.*  $n = 226$ ; *Est.* = Estimate of regression coefficient, *SE* = standard error, treatment is dummy-coded 0 = original, 1 = reconstructed,  $\Delta R^2$  = difference compared with a model without mediators. Statistically significant coefficients at the  $p \leq .05$  level are printed in bold.

**Table 2.3**

*Treatment Effects on Mediators (Facets of Perceived Authenticity) as well as Direct, Indirect, and Mediator Effects for Comparisons of the Original and Reconstructed Conditions*

	Cognitive processing and involvement						Physical processing and involvement			Learning outcomes		
	Vivid Imagination			Contextualization			Sensual Perception			Situational Interest		
	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>
<i>Mediator effects on dependent variables</i>												
Trust ( $b_1$ )	0.01	0.09	.886	-0.02	0.11	.816	0.11	0.10	.298	-0.05	0.08	.538
Typicality ( $b_2$ )	-0.04	0.06	.455	0.03	0.06	.614	0.06	0.07	.454	-0.10	0.06	.102
History ( $b_3$ )	<b>0.60</b>	<b>0.10</b>	<b>&lt;.001</b>	<b>0.38</b>	<b>0.11</b>	<b>.001</b>	<b>0.43</b>	<b>0.11</b>	<b>&lt;.001</b>	<b>0.57</b>	<b>0.08</b>	<b>&lt;.001</b>
Manipulation ( $b_4$ )	-0.04	0.05	.387	-0.06	0.06	.345	0.04	0.07	.572	0.00	0.05	.966
<i>Direct effects</i>												
Treatment ( $c'$ )	-0.15	0.12	.177	-0.102	0.13	.430	<b>-0.30</b>	<b>0.14</b>	<b>.035</b>	-0.07	0.11	.540
<i>Indirect effects</i>												
$a_1b_1$	-0.01	0.05	.891	0.01	0.05	.824	-0.05	0.05	.349	0.02	0.04	.559
$a_2b_2$	0.00	0.01	.869	0.00	0.01	.908	0.02	0.01	.871	0.00	0.02	.811
$a_3b_3$	-0.13	0.08	.120	-0.08	0.06	.176	-0.09	0.06	.149	-0.12	0.08	.113
$a_4b_4$	-0.02	0.03	.413	-0.03	0.03	.359	0.02	0.04	.595	0.00	0.03	.968
$R^2$	0.41	0.06		0.37	0.06		0.38	0.08		0.51	0.05	
$\Delta R^2$	0.31			0.13			0.30			0.23		

*Note.*  $n = 226$ ; *Est.* = Estimate of regression coefficient, *SE* = standard error, treatment is dummy-coded 0 = original, 1 = reconstructed,  $\Delta R^2$  = difference compared to a model without mediators. Statistically significant coefficients at the  $p \leq .05$  level are printed in bold.



### 3. Robustness Checks

We investigated how much the results reported for RQ1 or RQ2 were affected by differences across the two historical topics covered by the video clips and/or the six eyewitness accounts by looking at the adjusted mean differences for the original and reenacted conditions as well as the original and reconstructed conditions across historical topics and across the six video clips.

With regard to the historical topic, adjusted mean differences of all dependent variables (Situational Processing and Involvement, Sense of Insight, Situational Interest, Critical Distance) indicated that participants' ratings of these variables were higher in the original condition than in the reenacted and reconstructed conditions. In some cases (14 out of 60; the number of tests that resulted from testing two differences [original – reenacted; original – reconstructed] for 15 dependent variables for two historical topics [GDR and WWII]), results showed differences in the opposite direction, but these differences did not turn out to be statistically significant. For a comparison of the original and reenacted conditions, only one test of differences out of 30 was statistically significant (Unity of the Moment for the topic of WWII,  $\Delta M = 0.77$ ,  $SE = 0.22$ ,  $p < .001$ ). For a comparison of the original and reconstructed conditions, 11 out of 15 difference tests for the topic of WWII were statistically significant, ranging from  $0.23 < \Delta M < 0.89$ , all  $ps < .043$ ; whereas for the topic of GDR, only one difference between the original and reconstructed conditions turned out to be statistically significant (Unity of the Moment:  $\Delta M = 0.56$ ,  $SE = 0.22$ ,  $p = .010$ ).

Looking at the single video clips, mean differences indicated higher ratings for the variables in the original condition than in the reenacted or reconstructed conditions in most of the video clips and for most of the dependent variables. In the cases in which the result pattern was turned around, the mean differences were not statistically significant, with two exceptions (out of 180 tests; the number of tests that resulted from testing two differences [original – reenacted; original – reconstructed] for 15 dependent variables for six different video clips). However, due to the large number of tests, these results should be interpreted with caution as they might be due to chance.

Furthermore, we investigated the variance explained by models that included Treatment  $\times$  Clip or Treatment  $\times$  Topic interactions, respectively, in comparison with models without these interactions. Results are presented in Tables 3.1 and 3.2. The proportion of variance additionally explained by the Treatment  $\times$  Topic interaction ( $\Delta R^2$ , M1 – M0) compared with the variance additionally explained by the Treatment  $\times$  Clip interaction ( $\Delta R^2$ , M2 – M0) differed across all

the dependent variables, ranging from 0% (Perspective Recognition, original vs. reenacted) to 76% (Critical Distance, original vs. reconstructed). These results indicated that on average, 29% of the variance explained by the single video clips could be traced back to the topic of the video clips in the models comparing the original and reenacted conditions; and 33% of the variance explained by the single video clips could be traced back to the topic of the video clips in the models comparing the original and reconstructed conditions.

**Table 3.1***Comparison of Explained Variance in Models With Interaction Effects for the Original Versus the Reenacted Conditions*

	<b>M0</b>	<b>M1</b>	<b>M2</b>	<b>M2 vs. M1</b>	<b>M1 vs. M0</b>	<b>M2 vs. M0</b>	<b>Proportion</b>
	$R^2$	$R^2$	$R^2$				
	Covariates + Treatment only	Treatment × Content Interactions	Treatment × Clip Interactions	$\Delta R^2$ (M2-M1)	$\Delta R^2$ (M1-M0)	$\Delta R^2$ (M2-M0)	$\Delta R^2$ (M1-M0)/ $\Delta R^2$ (M2-M0)
Attentional Focus	0.07	0.09	0.17	0.07	0.03	0.10	0.26
Vivid Imagination	0.13	0.15	0.23	0.07	0.02	0.10	0.24
Abstract Imagination	0.20	0.21	0.28	0.07	0.01	0.08	0.15
Perspective Recognition	0.11	0.12	0.20	0.08	0.01	0.09	0.16
Contextualization	0.24	0.24	0.28	0.05	0.00	0.05	0.00
Being Moved	0.17	0.29	0.39	0.10	0.12	0.22	0.55
Immediacy	0.21	0.22	0.32	0.10	0.00	0.11	0.03
Unity of the Moment	0.18	0.25	0.33	0.08	0.07	0.15	0.46
Empathy	0.14	0.26	0.38	0.12	0.12	0.24	0.51
Similarities	0.09	0.14	0.21	0.07	0.05	0.12	0.43
Sensual Perception	0.09	0.10	0.15	0.05	0.01	0.06	0.12
Physiological Response	0.15	0.21	0.23	0.03	0.06	0.09	0.68
Situational Interest	0.24	0.24	0.33	0.09	0.00	0.09	0.01
Sense of Insight	0.11	0.13	0.19	0.06	0.02	0.08	0.22
Critical Distance	0.10	0.15	0.19	0.04	0.05	0.09	0.55

*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reenacted}} = 110$ . M0 refers to a model with the treatment predictor and all covariates; M1 adds a Treatment × Topic interaction term to M0; M2 adds Treatment × Clip interaction terms to M0.

**Table 3.2***Comparison of Explained Variance in Models With Interaction Effects for the Original Versus the Reconstructed Conditions*

	<b>M0</b>	<b>M1</b>	<b>M2</b>	<b>M2 vs. M1</b>	<b>M1 vs. M0</b>	<b>M2 vs. M0</b>	<b>Proportion</b>
	$R^2$	$R^2$	$R^2$	$\Delta R^2$	$\Delta R^2$	$\Delta R^2$	$\Delta R^2 (M1-M0)/$ $\Delta R^2 (M2-M0)$
	Covariates + Treatment only	Treatment $\times$ Content Interactions	Treatment $\times$ Clip Interactions	(M2-M1)	(M1-M0)	(M2-M0)	
Attentional Focus	0.05	0.07	0.17	0.10	0.02	0.12	0.20
Vivid Imagination	0.10	0.12	0.17	0.05	0.02	0.07	0.26
Abstract Imagination	0.10	0.15	0.20	0.05	0.04	0.10	0.46
Perspective Recognition	0.07	0.11	0.20	0.09	0.04	0.13	0.30
Contextualization	0.24	0.27	0.34	0.08	0.03	0.11	0.28
Being Moved	0.14	0.24	0.38	0.14	0.10	0.24	0.42
Immediacy	0.15	0.15	0.24	0.09	0.01	0.09	0.07
Unity of the Moment	0.21	0.29	0.35	0.07	0.08	0.14	0.53
Empathy	0.11	0.19	0.34	0.15	0.09	0.23	0.37
Similarities	0.06	0.10	0.15	0.06	0.03	0.09	0.37
Sensual Perception	0.08	0.09	0.15	0.06	0.00	0.07	0.06
Physiological Response	0.10	0.19	0.28	0.09	0.09	0.18	0.49
Situational Interest	0.28	0.29	0.34	0.05	0.01	0.06	0.10
Sense of Insight	0.08	0.10	0.17	0.06	0.02	0.08	0.23
Critical Distance	0.06	0.18	0.22	0.04	0.12	0.16	0.76

*Note.*  $n_{\text{original}} = 120$ ;  $n_{\text{reconstructed}} = 106$ . M0 refers to a model with the treatment predictor and all covariates; M1 adds a Treatment  $\times$  Topic interaction term to M0; M2 adds Treatment  $\times$  Clip interaction terms to M0.

---

# 6

## STUDIE 4

---

Zachrich, L., Wagner, W., Bertram, B., Nagengast, B. & Trautwein, U. (2022). When some students are better off with lessons “as usual”: An aptitude-treatment interaction study in the history classroom. *Zur Einreichung vorgesehenes Manuskript.*

**Abstract**

As postulated in research on aptitude-treatment interactions (ATIs), the effectiveness of an instructional setting can vary substantially for students with different characteristics. In this study, we investigated the interaction between students' dispositional comfort in social situations and the treatment effect of a socially demanding compared with a "standard" teaching unit (interviewing live eyewitnesses vs. texts/videos of eyewitness accounts in the history classroom, respectively) in predicting students' historical thinking competencies. Furthermore, we explored whether behavioral engagement mediated these interaction mechanisms. Data from a randomized controlled field trial on the effectiveness of an oral history intervention ( $N = 769$ ) were reanalyzed to examine Comfort  $\times$  Treatment interactions using multiple-group models in a structural equation modeling framework. Analyses revealed that students' comfort in social situations moderated the effect of the treatment on historical thinking competencies and that, in particular, these interactions were especially driven by students with low comfort benefiting more in socially less demanding settings. We did not find moderated mediation effects. We explore explanations and implications of the results, point out limitations and future research directions within the aptitude-treatment interaction framework.

*Keywords:* aptitude-treatment interaction, students' characteristics, behavioral engagement, mediation, history classroom

## Introduction

Both instructional designs and students' individual characteristics are key determinants in predicting students' achievement outcomes and learning (e.g., Hattie, 2009; Linnenbrink & Pintrich, 2003; Marzano et al., 2001). Research based on the *Aptitude-Treatment Interaction* (ATI; Cronbach, 1957) *paradigm* stresses the importance of students' characteristics (aptitudes) when examining the effectiveness of instructional methods (treatment) for a given outcome (Snow & Lohman, 1984; Yeh, 2012). In aptitude theory, aptitude is conceptualized as a student's readiness for a situation, depending on cognitive, conative, and affective characteristics and their combination (Ackerman, 2003; Snow, 1992). Initially beginning in the late 1950s, the quest to understand the variation in treatment effects for students with different learning prerequisites has been met with varying interest from educational psychologists over the decades; however, it is still an urgent issue today—and perhaps even more pressing than ever—as Reardon and Stuart (2017) stated in their Editors' introduction in the *Journal of Research on Educational Effectiveness*: “We need to develop and test theories that identify *how* and *for whom* and *under what circumstances* schooling and related interventions shape children's learning and development” (p. 671).

Classrooms are socially complex settings that require students to embody, among other skills, social and emotional competencies for handling daily classroom interactions successfully (Durlak et al., 2011; Frenzel et al., 2018). In history lessons in school, a social setting that is particularly demanding is considered a powerful instructional tool for enhancing learning: inviting eyewitnesses of the past to talk about their experiences (Busby & Hubbard, 2007)—thus interacting with a new and unacquainted person. This teaching format, in which guests are invited to attend history lessons to talk about their past, is generally believed to foster students' motivation to learn about history (Bertram et al., 2017) and engage with the teaching material (e.g., Patrick et al., 2007). However, Bertram et al. (2017) also found that students working with live eyewitnesses did not process the presented information as critically as students working with text or video versions of eyewitness accounts (indicated by two of three outcome measures at posttest). Among others, these results may be explained by the socially, emotionally, and cognitively demanding situation the live encounter evokes. Because of this socially demanding instructional setting, the Bertram et al. study seemed ideal for probing for ATI effects with affective moderators. Therefore, we reanalyzed the data to address the research question of how oral history interventions affect students' historical thinking competencies depending on students' readiness to handle these situations. Furthermore, we were interested in

the mediating role of self-reported behavioral engagement in the teaching unit within this aptitude-treatment framework.

### **ATI Research: Contributions and Limitations**

Everyone probably remembers enjoying a certain instructional setting the most in school, even when one's classmates did not engage or learn as much in these lessons but instead flourished in others. Thus, the effect of the instructional setting on learning success and engagement may depend on students' interest, motivation, or ability. Research dedicated to the question of the effectiveness of an instructional design on learning and achievement for students who differ in certain characteristics is at the heart of the ATI paradigm (Cronbach, 1957). The initial hypothesis of ATI research was that students with high ability profit in less structured classroom settings, whereas students with low ability need more structured guidance to profit from a classroom setting (Cronbach & Webb, 1975; Snow, 1977; Snow & Lohman, 1984). However, ATI research has not been restricted to ability. Moreover, whereas many researchers conceptualized aptitudes as fixed cognitive traits, others conceived aptitudes as malleable and situation dependent (Stemler & Sternberg, 2013).

Over the years, empirical support was found for the initial hypothesis and similar predictions. However, reviewing two decades of ATI research, Snow and Cronbach (1977) came to the disillusioning conclusion that replications of documented effects are rarely successful. As Cronbach and colleagues acknowledged, the main problems involved the complexity of the interactions. For example, many studies looked at a single aptitude of a group of students rather than "aptitude complexes" or suffered from a lack of explanatory power by only looking at pretest and posttest scores but not at the processes that take place in between (Ackerman, 2003; Snow, 1977; Snow et al., 2002). In addition, Snow (1989) identified the absence of social aspects of learning as a shortcoming of ATI research. Relatedly, statistical tests of interaction effects have less statistical power than tests of main effects (e.g., Spybrook et al., 2016), leading to an overestimation of statistically significant findings and difficulties in replication efforts (Gelman & Carlin, 2014). As a consequence, during the 1980s and 1990s, these difficulties led to an almost complete abandonment of what originally began as a highly intensive research field (Hedges & Schauer, 2018).

However, over the decades that followed, interest in heterogeneity within classrooms has grown again and has been the focus of different branches of research. Studies on learning styles and the so-called meshing hypothesis have proposed that learning is promoted when instruction matches the preferences of the learner (e.g., Kieft et al., 2008). Thus, a match



---

between preference and instruction results in optimal learning (Furnham, 2012). However, there is a lack of support for the existence and the effects of learning styles on competencies (Furnham, 2012; Pashler et al., 2009). A more successful approach for addressing heterogeneity within classrooms has been applied in research on adaptive teaching. This area of research emphasizes the idea that teachers need to consider students' individual strengths and prerequisites in order to foster their learning. Studies based on the micro-level of students' heterogeneity have tended to explore teachers' responses to students' learning, interest, or behavior (Corno, 2008; Paas et al., 2012; Parsons et al., 2017). On a macro-level, studies examining instructional or policy interventions have gone beyond looking at "only" the average treatment effect with the goal of creating a deeper understanding of the differential mechanisms of how educational practices affect students' learning (Corno, 2008; Reardon & Stuart, 2017).

Focusing on the macro-level, we found that researchers who have examined moderator effects of interventions have rarely used the idea of ATI to frame their analyses and that the studies that have referred to the ATI framework have frequently probed for cognitive aptitudes. For example, in an experimental setting, Fuchs et al. (2014) examined whether the effects of conceptual versus fluency activities on fourth graders' fraction knowledge were moderated by individual working memory. Results showed that students with weak working memory learned better with conceptual activities, whereas students with stronger working memory learned better with fluency activities (Fuchs et al., 2014). Fyfe and Rittle-Johnson (2016) examined whether the effects of a feedback intervention on students' math learning were moderated by students' prior knowledge and found that feedback helped students with low prior knowledge, but students with high prior knowledge profited more when there was no feedback. Recently, a special issue of the journal *Exceptional Children* focused on moderator effects in intervention research. All seven of the studies included in the special issue probed for initial skill levels (i.e., cognitive moderators) as moderators of the responsiveness to the intervention (Fuchs & Fuchs, 2019).

Conversely, very few studies have examined noncognitive factors as moderators of instructional treatments. One example is a recently published study by Flunger et al. (2018), who examined the roles of conative, cognitive, and affective moderators in an autonomy-supportive intervention in the physics classroom with respect to several outcomes. The authors explored whether gender, students' physics grade, and perceived autonomy moderated the effect of the intervention on different outcomes.

In sum, Snow (1992) originally conceptualized aptitudes as students' readiness to respond to a certain kind of instruction, comprising more than just cognitive competencies, and there has been plenty of anecdotal and scientific evidence for the roles that social and motivational factors play in learning (Linnenbrink & Pintrich, 2002, 2003; Mega et al., 2014; Zimmerman et al., 1992). Thus, it is somewhat surprising that there is a dearth of studies that have focused on ATI effects with regard to students' conative, affective, and social characteristics. Consequently, Reardon and Stuart (2017) called for differential examinations of treatment effects when they stated, "The field is now ripe for an expansion of focus" (p. 673). Hence, it has become clear that further research is needed to expand research on the interaction of students' social and motivational characteristics and instructional settings in predicting academic achievement in order to contribute to a more sophisticated picture of the adequateness of instructional settings for specific groups of students. An ideal setting in which to probe for such social and affective ATI effects is the setting of oral history interviews in history lessons because they include intense social interactions, and students with different levels of dispositional comfort in social situations are likely to handle these interactions differently.

### **Oral History Interviews as a Unique Instructional Setting**

In history lessons in school, inviting eyewitnesses to class is a popular method that has been recommended in books and handouts on history teaching (Oral History Association, 2009; Trškan, 2016; Whitman, 2004). Eyewitnesses of the past are people who witnessed a historical event or a certain time and share their emotional and individual perspective on their experiences. These experiences are often about abuse, neglect, dictatorship, and war (Trofanenko, 2017). In the *oral history* approach (e.g., Whitman, 2004), students are supposed to interview eyewitnesses of the past about their experiences, and this process is supposed to foster, among other benefits, students' historical thinking competencies.

The history teaching and history science communities have discussed several benefits (e.g., motivational potential) and risks (e.g., being overwhelmed by the emotional story) of the use of oral histories in school (Lanman & Wendling, 2006; Oral History Association, 2009; Trškan, 2016; Whitman, 2004). For example, students tend to report more interest in talking to an eyewitness in person because it is a more immediate, vivid, and concrete experience (Sochatzy, 2010). But they tend to trust eyewitness accounts more than text books and teachers (Angvik & von Borries, 1997; Rosenzweig & Thelen, 1998) and therefore have trouble critically reflecting on the eyewitness' story (Bertram et al., 2017; Hogervorst, 2018). In their intervention study, Bertram et al. (2017) examined the effectiveness of an oral history

---

intervention in school on students' historical thinking competencies with three experimental conditions. The intervention consisted of seven history lessons on the German Democratic Republic (GDR; one lesson = 45 min) that were identical across all intervention groups with the exception of two of the seven lessons that varied with regard to the eyewitness format (working with a live eyewitness, a video version of the eyewitness account, or a text version). Here, the eyewitnesses participated in the opposition movement in the GDR and shared their individual views on their role in the Peaceful Revolution. The results of the study indicated that students in the condition with a live eyewitness reported greater interest in learning with oral history interviews but did not profit as much as students in the video and text conditions with regard to their historical thinking competencies on two (understanding deconstruction, understanding oral history) of the three posttest measures. Whereas Bertram and colleagues examined the main effects of the eyewitness format on historical thinking competencies, they did not focus on differential effects of the intervention. However, moderation analyses have the potential to lead to a better understanding of the effectiveness of the teaching format for different students.

The encounter with a live eyewitness of the past offers a socially demanding setting for students—"it is more than the telling of experience but situated within a web of social processes" (Trofanenko, 2017, p. 154). Interviewing the eyewitness is a difficult task for students, as a teacher reported, "I think the hardest part on my end was helping them figure out what to ask, how to say it, order of the questions, how to handle responses, how to follow up responses, how much talking versus listening to do" (Teacher in Dutt-Doner et al., 2016, p. 262). Trofanenko (2017) even stated that students tend to have a certain level of anxiety that is associated with asking questions because the questions might evoke certain emotions—in themselves and in the witness. Overall, Yow (1997) argued that within this dynamic interview setting, a relationship between the interviewee and the interviewer evolves. This relationship might complicate the student's ability to learn to critically reflect on the story as part of historical thinking competencies, whereas watching a video or reading a text provides separation between the witness and the student and allows the student to take a more critical stance toward the source (Bertram et al., 2017; Hogervorst, 2018). Because social interactions in the classroom (e.g., interacting with the teacher or in classroom discussions) are not easy for some types of students (e.g., introverted students; e.g., Caspi et al., 2006; Medaille & Usinger, 2019; Murphy et al., 2017), interacting with an unacquainted other might be a social situation that is even more difficult to actively participate in. We argue that classrooms in which students are confronted with the personal and authentic story of an unacquainted other—in which they

are required to work with a dynamic source, ask questions, and interact with the new person—are socially demanding for some students. How students ask and interpret, respond to the narrator, and evaluate depends on their involvement with the person (Yow, 1997). Therefore, students are likely to handle this demanding situation differently depending on their characteristics and learning prerequisites. An oral history lesson may therefore offer a unique instructional setting for the investigation of ATI effects of students' noncognitive aptitudes and the instructional format on learning.

### **The Role of Student Characteristics in Learning in an Oral History Intervention**

How students perceive and handle academic and socially demanding situations depends, among other qualities, on their characteristics such as social emotional skills, motivation, or personality (DiPerna et al., 2018; Durlak et al., 2011; Flunger et al., 2018; Muenks et al., 2017). The oral history classroom with live eyewitnesses presents students with an unacquainted person and an individual and authentic story about the past. Shy students can be expected to withdraw from this social interaction with the eyewitness, whereas students who feel comfortable in social situations are more likely to more actively engage and participate in the dynamic classroom conversation. Therefore, comfort in social situations might be likely to influence how students engage and learn in the oral history setting.

Literature on the kinds of characteristics that affect students' handling of social situations emphasizes the role that a student's level of comfort in these new situations plays. For example, shyness is conceptualized as wariness, fear, and discomfort in social situations. Korem (2018) stated that sensitivity to novelty is one crucial inhibition of shy people and that it primarily manifests as fear and hesitancy when encountering new people (Kagan, 1997). Similarly, students who were not well-accepted by their classmates were found to avoid new and complex social situations (Jerusalem & Klein-Hessling, 2009). In addition, social self-efficacy—the feeling of being able to cope with new and difficult social situations—has been found to be associated with competent social behavior in socially demanding situations (Connolly, 1989; Jerusalem & Klein-Hessling, 2009). Bilgin (1996) suggested that higher levels of social self-efficacy help students to be more confident when engaging in interpersonal relations (see also Tras, 2016). Finally, high self-esteem was found to affect confident behavior when meeting unacquainted others (Krause et al., 2016). Putting all this evidence together, we argue that these variables contribute to students' dispositional comfort in social situations. In particular, these variables can be expected to affect how students will handle an encounter with an eyewitness of the past in the history classroom. Like other aptitude variables, we do not

---

consider dispositional comfort in social situations to be unchangeable over time, but rather—at least to a certain degree—being malleable.

The available studies suggest that students' comfort might not directly affect their competencies but might rather affect their engagement in the classroom, which can then be expected to affect their learning. Several empirical studies have found evidence for the importance of students' behavioral engagement for academic achievement: Students who actively participate in the classroom, work on the given tasks, and interact with teachers and peers have been found to achieve better learning outcomes (e.g., Fredricks et al., 2004; Nguyen et al., 2018). The association between students' comfort in social situations and their competencies can therefore be expected to be mediated by their engagement in an instructional setting in which they meet and interact with an eyewitness of the past.

### **Research Questions**

We aimed to address two crucial shortcomings of ATI research with the present study: the first involving affective aptitude variables and the second involving the processes that take place between pretest and posttest through a mediator (Snow et al., 2002). Drawing on data from the randomized controlled field trial by Bertram et al. (2017) with several motivational and affective predictors enabled us to address the following research questions: Do the effects of a socially demanding instructional setting (as compared with an instructional setting “as usual”) on students' competencies differ for students with different levels of dispositional comfort in social situations (Research Question 1)? Does students' dispositional comfort in social situations have an effect on their behavioral engagement in the different instructional settings (Research Question 2)? And if so, does behavioral engagement mediate the effect of the interaction between students' dispositional comfort in social situations and the instructional setting on students' competencies (Research Question 3)? We expected that an encounter with a live eyewitness would be especially likely to be associated with students' historical thinking competencies rather than their knowledge because eyewitnesses talk more about personal experiences and perceptions than about facts and figures. Additionally, incorrect facts would be corrected by the teacher.

Because we considered different effects to be plausible (e.g., students with high comfort profiting in the live condition but also students with low comfort profiting in the video/text condition), we approached our research questions in a rather exploratory manner without specific hypotheses. Although we are referring to moderation—or moderated mediation—

*effects*, these effects are not to be interpreted as causal because the interaction term only represents variation in a given (causal) effect across levels of the moderator (Edwards, 2009).

### **Method**

The present investigation involved a reanalysis of data from a randomized controlled field trial on the effectiveness of an oral history intervention in history lessons in German schools conducted by Bertram et al. (2017). The study was funded by the Federal Ministry of Education and Research, Germany (Project Number 01JG0913). By the time the study took place, the local educational authorities or the state ministry of education were required to evaluate empirical studies in terms of ethical issues and questions pertaining to data security. The study was approved by the local authorities of Tübingen (Reference Number: 72-3/6499.20). Active consent was collected from the parents of all participating students.

The original study examined the effectiveness of a seven-lesson teaching unit with regard to fostering students' historical thinking competencies in three experimental conditions (a live eyewitness in class, a video version of the eyewitness account, or a text version) and an untreated control condition. The data were collected within a pre-post-follow-up test design where the pretest ( $t_1$ ) was assessed immediately before the teaching unit began, the posttest ( $t_2$ ) was assessed in the first regular history lesson after the teaching unit (after the evaluation of the encounter with the eyewitness account), and the follow-up test ( $t_3$ ) was administered around 3 months after the teaching unit. For a detailed description of the study, see Bertram et al. (2017).

### **Sample**

For the current analyses, we excluded the untreated control condition because these classes did not explore any eyewitness accounts at all. Every treatment condition consisted of 10 classes each: live condition:  $N = 250$  students; video condition:  $N = 256$  students; text condition:  $N = 263$  students. The total sample consisted of  $N = 769$  students in 30 classes and three treatment conditions. Participation rates were on average 98.22% in the classes experiencing the live condition, 96.12% for the video condition, and 100% for the text condition. The average age of the students was 14.52 years ( $SD = 0.61$ ), and about 45% were female adolescents. The participating classes belonged to the higher track secondary schools in the federal state of Baden-Württemberg in Germany.

On the basis of the current sample, we calculated post hoc power analyses to find the minimum detectable effect size difference (MDESD) of the individual-level moderator effects (Spybrook et al., 2016) for a statistical power of .80. We used manifest regression models for the calculations and expected that the results would be comparable to our latent regression

analyses. The calculations revealed detectable standardized effect size differences for all outcomes at posttest and follow up ( $0.17 \leq \text{MDESD} \leq 0.18$ ). The effect size difference is defined as the difference between the main effects of an intervention on a standardized outcome for two subgroups (Bloom & Spybrook, 2017).

## Measures

**Students' dispositional comfort in social situations.** Of the variables that were available in the data set, we used social self-concept, social self-efficacy, and self-esteem as indicators of students' dispositional comfort in social situations. Social self-concept, social self-efficacy, and self-esteem were each assessed with a 4-point Likert scale (1 = *I don't agree at all* to 4 = *I completely agree*) at the first measurement point. We provide the full list of items to these constructs in Appendix A.

**Social self-concept.** Social self-concept was assessed with a four-item scale that focused on questions about students' acceptance in class (Fend & Prester, 1986; Leung et al., 2015). A sample item is "Sometimes I feel a little bit like an outsider in my class" (translated, reverse coded). The scale contained three reverse-coded items and had good internal consistency (Cronbach's  $\alpha = .80$ ). The latent construct social self-concept, specified by four indicators, showed an adequate model fit,  $\chi^2(2) = 4.54, p = .103, \text{RMSEA} = 0.04, \text{CFI} = 0.99, \text{SRMR} = 0.01$ .

**Social self-efficacy.** The eight-item social self-efficacy scale was used to assess students' expectations of their ability to handle a demanding social situation and was adapted from Jerusalem and Klein-Hessling (2002). It included three dimensions of social self-efficacy: meeting new fellow students (3 items), handling conflict situations (3 items), and convincing others of one's own opinion (2 items). A sample item is "I can easily chat with fellow students, even if I don't know them" (translated). According to the developers of the scale, the total social self-efficacy scale is sufficiently homogenous to measure one construct (Jerusalem & Klein-Hessling, 2002). However, in the present study, a confirmatory factor analysis of a one-factor model showed an inadequate model fit,  $\chi^2(20) = 127.86, p < .001, \text{RMSEA} = 0.09, \text{CFI} = 0.84, \text{SRMR} = 0.05$ . However, a second-order factor model for the three theoretically assumed constructs was found to show an adequate model fit for social self-efficacy,  $\chi^2(17) = 33.15, p = .011, \text{RMSEA} = 0.04, \text{CFI} = 0.98; \text{SRMR} = 0.04$ . We therefore decided to model social self-efficacy as a latent variable specified by three manifest scales in accordance with the theoretical assumptions. The three manifest scales were moderately and statistically significantly correlated ( $.31 < r < .38$ ). Cronbach's  $\alpha$  in the current sample was  $\alpha = .71$ .

**Self-esteem.** Fend and Prester's (1986) global self-esteem scale was used to identify a student's overall picture of him- or herself. A sample item from the three-item scale is "Overall, I have a positive image of myself" (translated). The internal consistency of the scale was adequate (Cronbach's  $\alpha = .82$ ).

**Process variable: Behavioral engagement.** Using four posttest items administered after the teaching unit, students were asked how much effort they put into the tasks and how engaged they were in the teaching unit. Sample items for this construct are: "I worked as hard as I could during the teaching unit *Peaceful Revolution of the GDR*" and "I worked very accurately with the materials and the reader so I could fulfill all tasks" (translated). The items were self-developed by the authors who conducted the data collection. The internal consistency of the behavioral engagement scale was adequate (Cronbach's  $\alpha = .77$ ).

**Outcome variables: Historical thinking competencies.** In line with Bertram et al. (2017), we examined three outcome variables focusing on students' historical thinking competencies (see Bertram et al., 2017, for a detailed description of the teaching unit and the outcome variables). Insights into epistemological principles of history display one part of historical thinking competencies and a major goal of history education. In particular, three aspects of insights into epistemological principles of history were considered within this study. All measures were assessed at all three time points. The reliability calculations for the achievement test measures were based on the original sample of  $N = 962$ . All items were assessed with a 4-point Likert scale, ranging from 1 = *I am completely sure that the statement is not true* to 4 = *I am completely sure that the statement is true* at all measurement points.

**Understanding reconstruction.** The first measure assessed the insight that history has to be reconstructed on the basis of sources and accounts. A sample item from this measure is: "Historians have to take into account the possibility that sources might have gaps" (translated). This measure contained a total of 24 items and had a Weighted Likelihood Estimator-Person Separation Reliability (WLE PSR) = 0.87. The WLE PSR is a measure of reliability (Andrich, 1982) and is defined as 1 minus the relative amount of measurement error, with the relative amount of measurement error defined as the average squared standard error of each person's WLE (i.e., variance due to measurement error) divided by the variance of the WLEs (i.e., total variance due to variance in the person's "true" ability and measurement error).

**Understanding deconstruction.** The second measure assessed the insight that historical accounts have to be analyzed in detail. A sample item is: "If you read something in a history



text book, you can be sure that history has happened exactly as described in the text” (translated, reverse coded). The measure consisted of six items, WLE PSR = 0.59.

**Understanding oral history.** The last outcome measure focused on insights into the particularity of an eyewitness account and the person’s perspective on the historical event. A sample item from this 10-item measure is, “Every eyewitness has his/her own view on the past” (translated), WLE PSR = 0.73.

**Treatment.** In line with the procedure chosen by Bertram et al. (2017), students in the video and text conditions (“recorded oral histories” as compared to active oral history) were treated as one condition in the current study. We expected that working with recorded oral histories would reflect lessons as usual, whereas lessons with live eyewitnesses seemed to be extraordinary and more complex because students encountered actual unacquainted others. To check the decision to combine video and text condition for robustness, we ran the final analyses for live vs. video and for live vs. text. Results of the moderation models indicated the same results pattern for the aptitude-treatment interactions in both models: live vs. video and live vs. text. With regard to the mediation models, we also found the same results pattern of the interaction terms in both groups. However, they differed with regard to the size of the effect, whereas calculations for the live vs. video group revealed larger effects than calculations for the live vs. text group (analysis scripts with results are available at the Open Science Framework: <https://osf.io/9h7pz/?viewonly=739401c47f404791bdd47dc85fe325f1>). In order to keep enough power, we decided to stick with our decision and the procedure Bertram et al. (2017) chose to model these lessons with recorded oral histories as “lessons as usual.” Hence, the treatment variable, group membership in one of the two experimental conditions (video/text vs. live), was dummy-coded as 0 = video/text and 1 = live.

## Analysis

**Analytical models.** To examine our research question about the moderating effect of students’ comfort in social situations on the intervention effect (RQ 1), the effect of comfort in social situations on behavioral engagement (RQ 2), and the mediating effect of behavioral engagement within this model (RQ 3), we computed regression models in which we combined manifest variables (dependent variables and covariates) and latent variables (dispositional comfort and behavioral engagement). The moderator variable—dispositional comfort in social situations—was modeled as a second-order factor with the three latent first-order factors social self-concept, social self-efficacy, and self-esteem. Because we had a dichotomous treatment variable (live vs. video/text), we used multiple-group models with one group for each of the

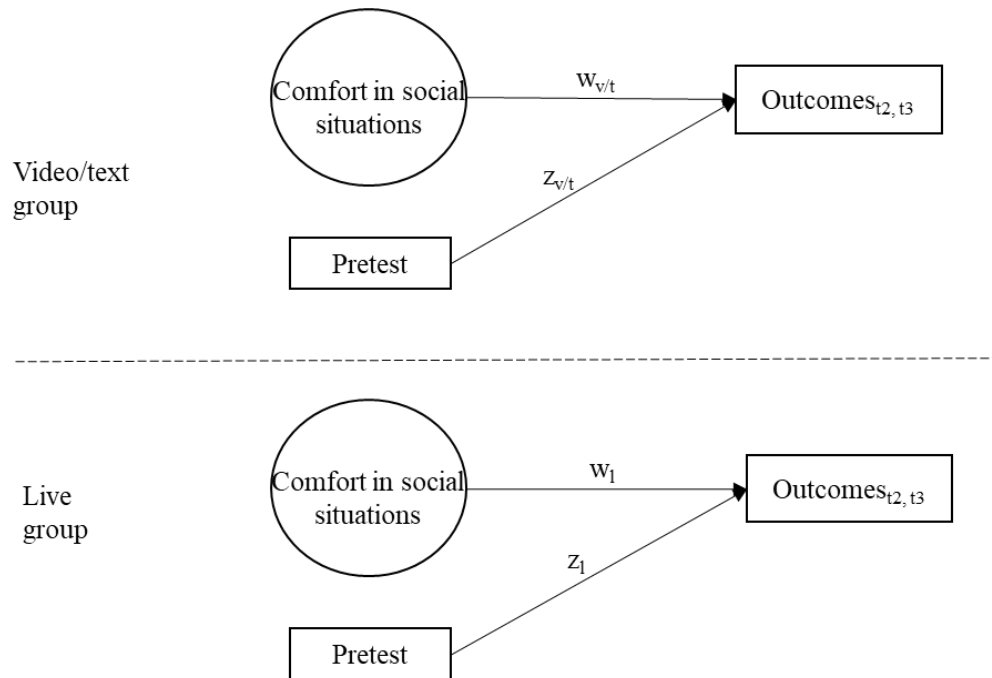
two conditions to predict students' historical thinking competencies through their comfort in social situations (moderation; Figure 1; Kline, 2016; Ryu & Cheong, 2017). We tested the conditional effect by estimating the difference between the regression coefficients in the two groups ( $w_1 - w_{v/t}$ ). Expanding this model for our second research question, we included behavioral engagement to mediate the effect of the interaction between comfort in social situations and the treatment on students' competencies (Figure 2). We were first interested in the (group-specific) effects of comfort on behavioral engagement ( $a_{v/t}$ ,  $a_1$ ) as a process measure. We then tested the mediation by investigating the effect of engagement on the competence measures ( $b_{v/t}$ ,  $b_1$ ), the direct effect of comfort on the competence measures ( $c_{v/t}$ ,  $c_1$ ), and the indirect effects ( $ab_{v/t}$ ,  $ab_1$ ). We tested the conditional indirect effects by estimating the difference between the indirect effects for the two groups ( $ab_1 - ab_{v/t}$ ; first stage moderation; Figure 2; Edwards & Lambert, 2007). These differences were tested with the Model Constraint option in Mplus.

Before we estimated the regression models, we tested for measurement invariance for the latent variables (Chen et al., 2005). Results indicated that strong factorial invariance could be assumed (see Appendix B for a detailed report).

We began the examination by computing confirmatory factor analyses to see if latent factor models for moderators and mediators would show an adequate fit. Therefore, we first estimated the factor loadings for the second-order factor in a supplementary model without grouping, and we set the factor variance to 1 and factor mean to 0.

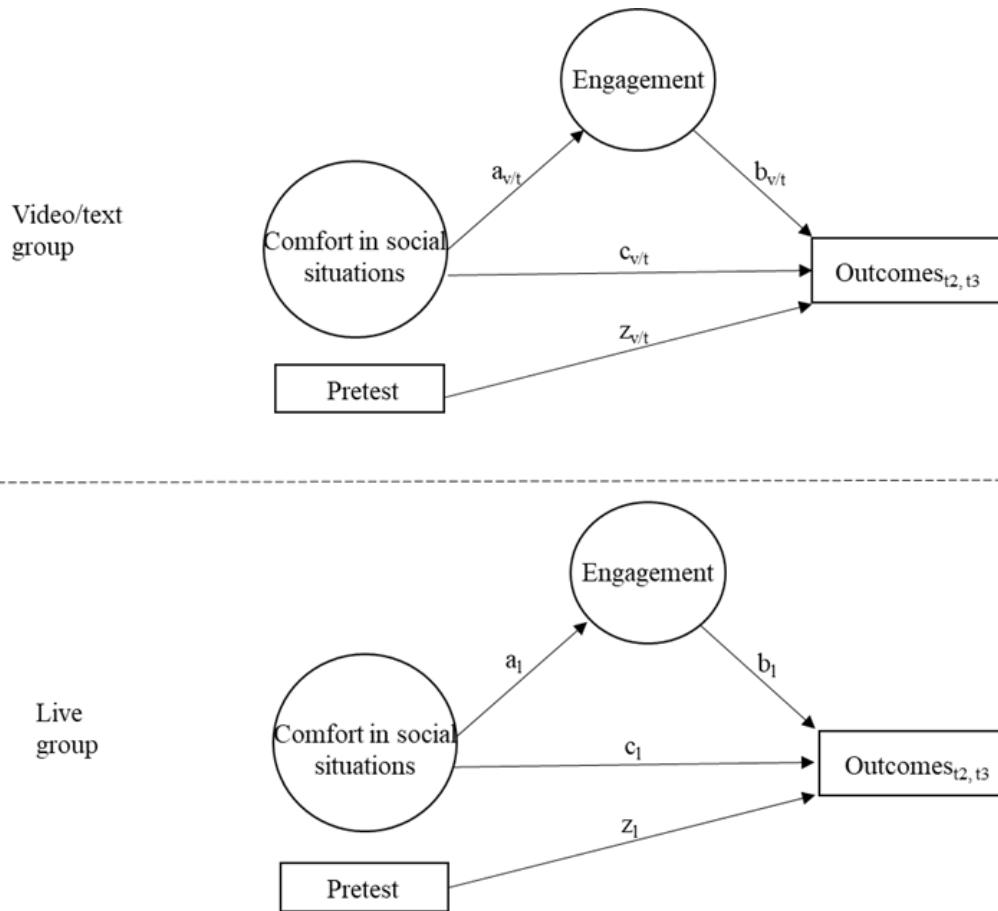
**Figure 1**

*Multiple-Group Model for the Regression of Students' Comfort in Social Situations on Insights Into Epistemological Principles at Posttest ( $t_2$ ) and Follow-up ( $t_3$ ) Controlled for All Pretest Scores ( $t_1$ ) in the Video/Text and Live Conditions*



**Figure 2**

*Multiple-Group Model for Students' Comfort in Social Situations on Insights Into Epistemological Principles via Behavioral Engagement Controlled for All Pretest Scores in the Video/Text and Live Conditions*



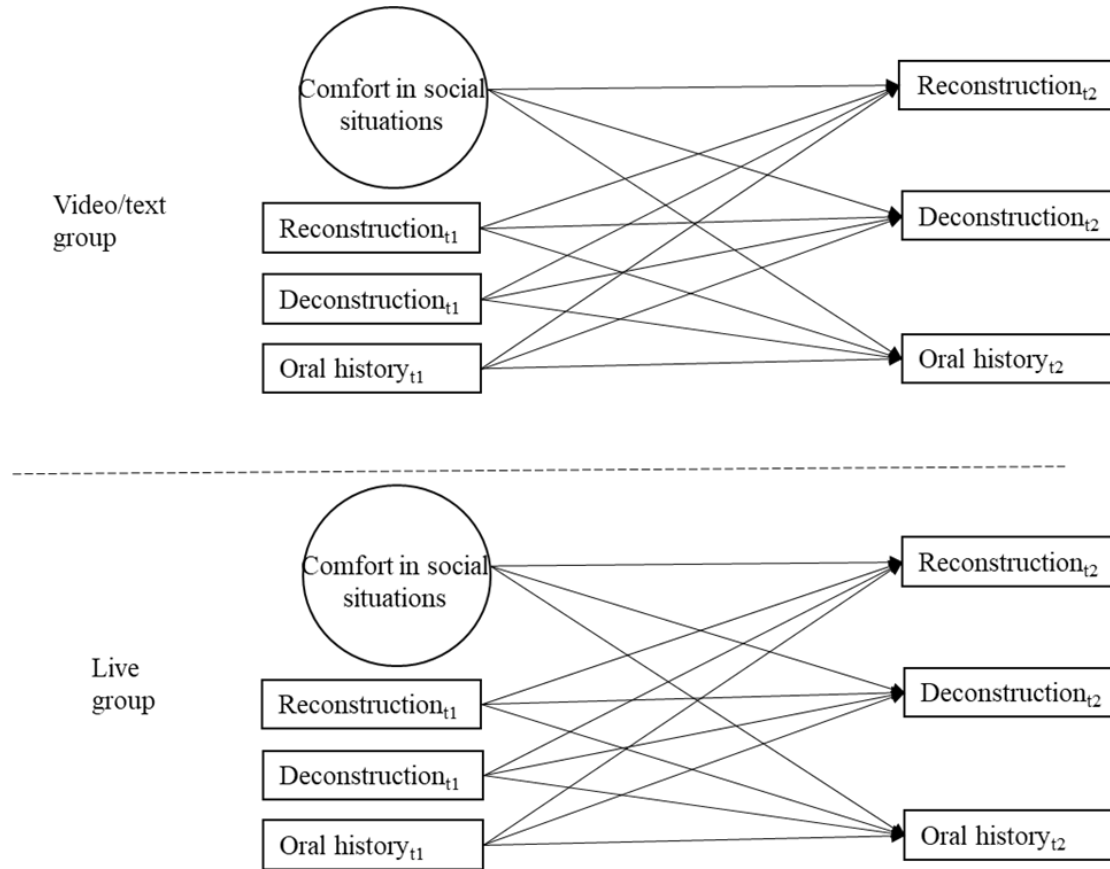
We specified our models by fixing the factor loading of the first indicator of the first-order factors according to the factor loading value we found in the confirmatory factor analyses without grouping (in which the overall variance of the latent variables had been fixed to 1 for identification), and we let the variance be estimated freely in both groups. By doing this, we allowed for an interpretation of the results in terms of effect sizes based on overall (i.e., across students in the whole sample) standardized factors—while accounting for the possibility that there would be different variances in the groups. The intercepts of the first-order factors were constrained to be zero in both groups, and the intercepts of the respective indicators were constrained to be equal across groups. We specified the second-order factor in the final models by constraining the mean to be equal to zero in the first group, but we freely estimated the mean in the second group (to allow for potential differences in factor means between the two groups).

Possible treatment-covariate interactions (i.e., different regression coefficients for the pretests on the posttest or follow-up tests) were tested using chi-square difference tests (adjusted by the scaling correction factors to avoid inflated overall goodness-of-fit test statistics; Bryant & Satorra, 2012) by holding these parameters equal across groups. We tested the specified model with group-specific regression coefficients against a model with equality constraints for all respective pretest-specific coefficients across groups in terms of an omnibus test for the moderation effect (e.g., all paths from the pretest scores to the outcome measures were held equal; Figure 3). Moreover, each single moderation effect was tested with a *z* test of the respective difference in parameters between the two groups (i.e., the effect of comfort in social situations on the dependent measures or the difference in the effect of behavioral engagement on the dependent measures). The differences in parameters were estimated using the Model Constraint option in Mplus (Muthén & Muthén, 1998-2017).

In total, we set up four multigroup models: two models for examining the research question about the moderation (Model 1: posttest; Model 2: follow-up) and two models for examining the effects on behavioral engagement and the moderated mediation (Model 3: posttest; Model 4: follow-up). Figure 3 shows Model 1 as an example. We estimated an additional model without grouping to investigate the global effect of engagement on the outcome measures.

**Figure 3**

*Example Model for Students' Comfort in Social Situations on All Three Outcome Measures at Posttest Controlled for Pretest Scores*



For the first research question (RQ 1), we tested for a difference in the regression coefficients for dispositional comfort in social situations on historical thinking competencies between groups. In our framework, this difference can be interpreted as the conditional effect of the treatment depending on students' comfort—or vice versa. To aid the interpretations of the interaction terms, we compared these differences for students with low comfort versus students with high comfort in social situations with regard to the outcome measures. By doing this, we gained insights into the treatment effects for the subgroups. We set up equations to predict students' outcomes in the live and video/text conditions for specific values of comfort in social situations (+1 *SD* and -1 *SD* referring to the total variance across the two groups) with the Model Constraint option in Mplus. The following equation displays the pattern for these analyses.

$$\hat{Y}_{ij-high} = w_{0ij} + z_{1j}\bar{Z}_1 + z_{2j}\bar{Z}_2 + z_{3j}\bar{Z}_3 + w_{1ij}(SD_{css})$$

$$\hat{Y}_{ij-low} = w_{0ij} + z_{1j}\bar{Z}_1 + z_{2j}\bar{Z}_2 + z_{3j}\bar{Z}_3 + w_{1ij}(-SD_{css})$$

The outcome  $Y_i$  in condition  $j$  for students with high comfort ( $\hat{Y}_{ij\text{-high}}$ ) was predicted through an outcome- and group-specific intercept ( $w_{0ij}$ ), the group-specific regression coefficients for each pretest measure ( $z_{1j}\text{-}z_{3j}$ ) multiplied by the overall mean of the respective covariate ( $\bar{Z}_1\text{-}\bar{Z}_3$ ), and the group-specific regression coefficient for students' comfort in social situations on the respective outcome multiplied by  $SD = 1$  for students with high comfort and by  $SD = -1$  for students with low dispositional comfort in social situations. We tested the different treatment effects for both subgroups for significance.

For our second research question (RQ 2), we tested the group-specific association between students' dispositional comfort in social situations and behavioral engagement in the teaching unit ( $a_l, a_{v/t}$ ; Figure 2). Again, we compared students with high and low comfort in social situations with regard to students' behavioral engagement to test the treatment effects for student subgroups. By doing so, we predicted the expected engagement for students with high (+1  $SD$ ) and low (-1  $SD$ ) comfort in the live and the video/text conditions, respectively. From these calculations, we generated a plot to evaluate the role of comfort in social situations in moderating the effect of the intervention on the three competence measures (i.e., the outcomes variables) and behavioral engagement.

Pertaining to the last research question (RQ 3), we tested the group-specific direct and indirect effects of comfort in social situations on the competence measures via behavioral engagement before we estimated the difference in the indirect effects between the live and video/text groups ( $ab_l - ab_{v/t}$ ; Figure 2). The results can be interpreted as standardized effect sizes (Wen & Fan, 2015).

All analyses were computed in Mplus (version 8; Muthén & Muthén, 1998-2017), and the clustering of the data was taken into account by using the design-based correction of standard errors and model fit statistics (Mplus command: Type = COMPLEX) by using robust maximum likelihood estimation. This approach lies on less assumptions than typical multilevel models, yields correct inferences taking the nested data structure into account and is appropriate in situations where variables at the cluster-level are not the main focus (McNeish, Stapleton, & Silverman, 2017). All data analysis scripts are available for review at the Open Science Framework at the following address: [https://osf.io/9h7pz/?view\\_only=739401c47f404791bdd47dc85fe325f1](https://osf.io/9h7pz/?view_only=739401c47f404791bdd47dc85fe325f1) (the project will be publicly available after the review process is finished).

**Missing data.** Although active consent was collected from all parents, students did not necessarily participate at all time points. Overall, the sample consisted of 769 students (519 students in the video/text condition, 250 in the live condition). In the video/text condition,

92.87% of the students participated at pretest, 93.83% participated at posttest, and 90.94% participated at follow-up. In the live condition, 92% of the 250 students participated at pretest, 91.6% participated at posttest, and 88.8% participated at follow-up. A total of 7%-12% of the responses across all variables used in the analyses were missing due to nonresponse or invalid responses. To handle missing data, we used a full information maximum likelihood estimator (FIML; Enders & Bandalos, 2001; Muthén & Muthén, 1998-2017).

**Model fit.** We evaluated the model fit by computing the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Standardized Root Mean Square Residual (SRMR). RMSEA values below 0.05, CFI values above 0.95, and SRMR values below 0.08 are considered indicative of a satisfactory to good model fit (Yu, 2002). The competence measures were standardized by using the mean and standard deviation of the respective pretest measure, so the results could be interpreted with regard to the metric of the pretest scores. There are two threats that can occur when statistical results are interpreted on the basis of  $p$ -values: (a) overlooking results that might be important and (b) interpreting results as important when they were actually random. Referring to the recent discussion on science without  $p$ -values (Wasserstein et al., 2019), we therefore tried to first present and interpret the results with regard to the pattern of the actual coefficients we found. However, to safeguard against finding only effects due to sampling error, we also controlled the false discovery rate (FDR) as proposed by Benjamini and Hochberg (1995). To define the families of statistical tests, we summarized the tests with regard to the sets of dependent variables separated by our research questions, thus resulting in 18 tests for the moderation models with competence measures as dependent variables (RQ 1, Table 3), three tests of the associations between comfort and engagement and their interaction (RQ 2;  $a_1$ ,  $a_{v/t}$ ,  $\Delta a$ ), and 18 tests for the moderated mediation models analyzing competence measures (RQ 3; Table 6). The remaining 18 tests for the direct paths of the mediation analysis were summarized into a fourth family.



## Results

### Descriptive Statistics, Correlations, and Preliminary Results

Descriptive statistics for the three constructs representing dispositional comfort in social situations as well as behavioral engagement are displayed in Table 1 for the two treatment conditions. The latent correlations of these constructs are displayed in Table 2. The comfort constructs were moderately to highly correlated ( $.42 \leq r \leq .74$ , all  $ps < .001$ ). The confirmatory factor analysis of a second-order factor model for dispositional comfort in social situations, with social self-concept, social self-efficacy, and self-esteem as the first-order factors, revealed an adequate fit,  $\chi^2(31) = 61.80$ ,  $p < .001$ ; RMSEA = 0.04; CFI = 0.98; SRMR = 0.03. Correlations between comfort in social situations and all outcome measures (i.e., the historical thinking competencies) at posttest and follow-up were small ( $-.07 \leq r \leq .13$ ). The outcome variables were moderately to strongly intercorrelated ( $.48 \leq r \leq .64$ ). The intercorrelations of behavioral engagement, the outcomes, and comfort were small in size ( $-.15 \leq r \leq .27$ ).

**Table 1**

*Descriptive Statistics for Comfort in Social Situations and Behavioral Engagement (Manifest Scales) in the Two Treatment Conditions*

	Live ( $N = 250$ )			Video/text ( $N = 519$ )		
	$N$	$M$	$SD$	$N$	$M$	$SD$
Social self-concept	214	3.35	0.54	446	3.35	0.54
Social self-efficacy	217	2.97	0.38	453	2.96	0.37
Self-esteem	223	3.17	0.64	465	3.22	0.56
Engagement	222	2.77	0.67	476	2.56	0.67

*Note.* Minimum = 1, Maximum = 4.

To ensure that the treatment-covariate interactions could be ignored, we computed chi-square difference tests for models with freely estimated covariates (less restrictive,  $M_0$ ) versus models with equality constraints between groups (more restrictive,  $M_1$ ). The analyses showed nonsignificant results for all models and therefore implied that treatment-covariate interactions could be ignored (i.e., effects for the covariates were held constant across groups; adjusted by the scaling correction factors,  $s$ , Bryant & Satorra, 2012); posttest:  $M_0: \chi^2(186) = 310.68$ ,  $s = 1.03$ ;  $M_1: \chi^2(195) = 316.73$ ,  $s = 1.04$ ,  $p = .831$ ; follow-up:  $M_0: \chi^2(186) = 306.06$ ,  $s = 1.03$ ;  $M_1: \chi^2(195) = 317.86$ ,  $s = 1.04$ ,  $p = .405$ . All models showed good fit.

**Table 2**

*Latent Correlations for Comfort in Social Situations, its Constructs, and Behavioral Engagement*

	1	2	3	4	5
1 Comfort in social situations	-				
2 Social self-concept	.81***	-			
3 Social self-efficacy	.88***	.71***	-		
4 Self-esteem	.55***	.45***	.49***	-	
5 Engagement	.10*	.09	.09*	.06	-

*Note.* Latent correlations were estimated in Mplus. Missing data were handled with the full information maximum likelihood estimator.  $N = 759$ . \*  $p < .05$ . \*\*\*  $p < .001$ .

### **Moderation of the Treatment Effect by Students' Comfort in Social Situations**

For our first research question (RQ 1), we tested for whether the treatment effect of the oral history intervention on historical thinking competencies was moderated by students' dispositional comfort in social situations. To this end, we calculated multigroup models for the outcome measures at posttest and follow-up (see Figure 1) and investigated group-specific regression coefficients as well as their interactions.

Results of the group-specific regression analyses and the interaction effects (Figure 1:  $w_1 - w_{v/t}$ ) are displayed in Table 3. At posttest, we found negative regression coefficients for comfort in predicting students' historical thinking competencies in the video/text condition, and positive regression coefficients in the live condition. The interaction terms were positive and ranged from  $\Delta w = 0.22$  (understanding reconstruction) to  $\Delta w = 0.29$  (understanding oral history). Overall, these results indicated a negative association of students' comfort in social situations and the competence measures in the video/text condition but a positive association in the live condition and positive interaction effects. The same pattern emerged for the measures at follow-up (negative association in the video/text condition, positive association in the live condition, except for the effect of comfort on deconstruction, positive interaction terms). Looking at the magnitudes of the standardized effects, the results indicated substantial moderator effects (estimates:  $0.12 \leq \Delta w \leq 0.41$ ) compared with the treatment effect values for the main effects (live vs. video/text; estimates:  $-0.23 \leq \text{Est.} \leq 0.02$ ; see Bertram et al., 2017, p. 26). Testing these coefficients for significance controlling for the FDR, we found 7 out of 18 results to be statistically significant.

---

Pertaining to our first research question, results indicated that it seems to be the negative impact of students' comfort in social situations on competencies in the video/text condition that is driving the moderation. While students feeling higher dispositional comfort in social situations seem to profit in both conditions comparably, students' with lower comfort in social situations benefit more in the video/text condition.

We additionally specified models with each of the first-order factors as moderators (social self-concept, social self-efficacy, and self-esteem) instead of using a second-order factor to see if the results with regard to dispositional comfort in social situations could be replicated for each of the three constructs. As reported in Appendix C, we found a similar pattern with respect to the point estimates (direction of the coefficients: negative coefficients in video/text condition; positive coefficients in live condition; positive interaction terms) although the pattern of statistical significance slightly changed.

To facilitate interpretation, we predicted competence differences in the two treatment conditions (the outcomes at posttest and follow-up compared with the pretest) for students with low ( $-1 SD$ ) and high ( $+1 SD$ ) comfort in social situations. Figure 4 displays the association between the treatment effect on the respective competence measure (y-axis) for students with low and high comfort (x-axis). All of the displayed plots suggest that students with low comfort in social situations profited more from the video/text treatment than from the live treatment, and conversely, students with high comfort in social situations profited more in the live condition than in the video/text condition, or in both conditions comparably. Nevertheless, the plots indicate that students' comfort in social situations moderates the treatment effect on historical thinking competencies.

**Table 3**

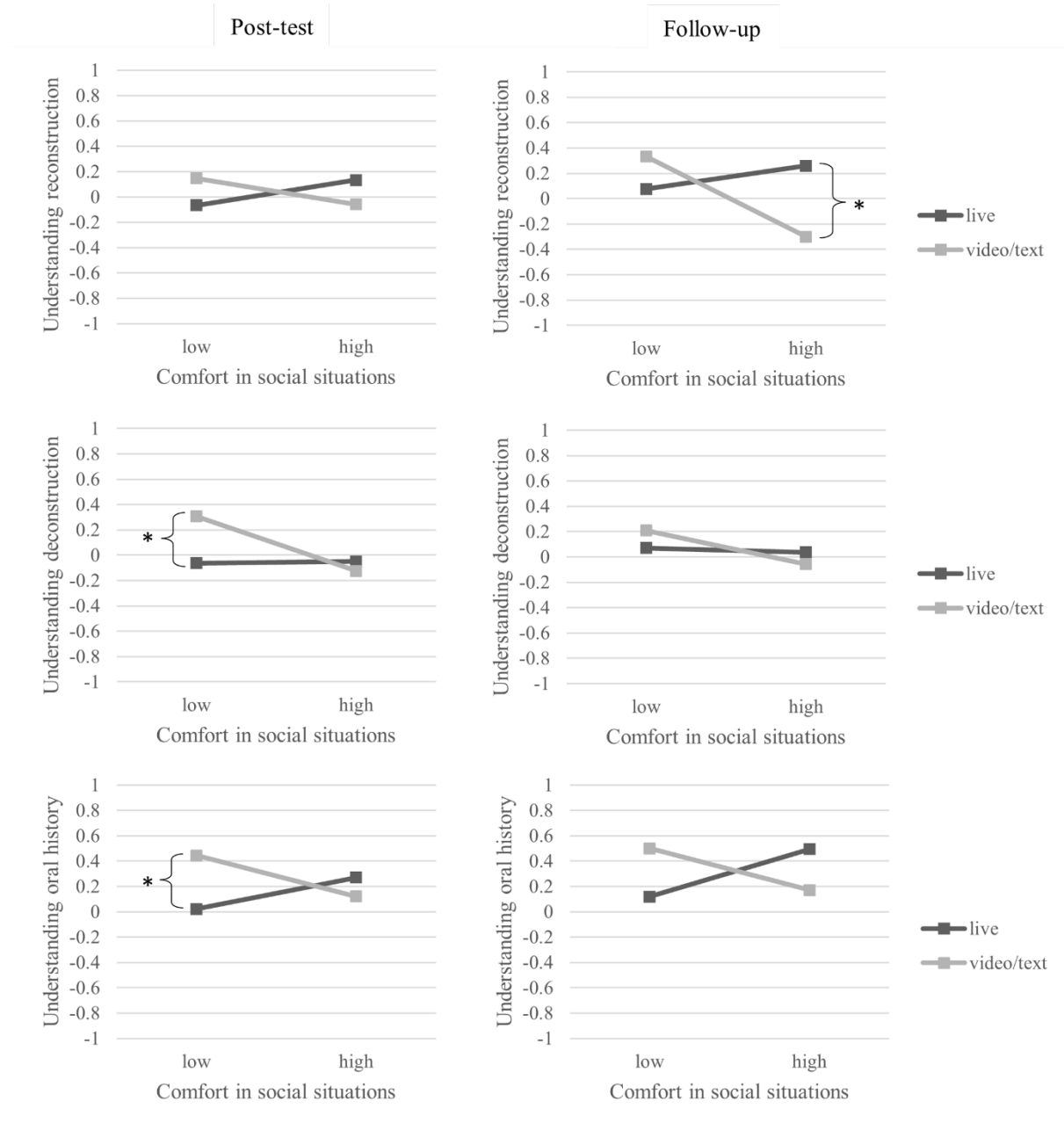
*Moderation Analysis: Regression Coefficients for Comfort in Social Situations in Predicting the Outcomes at Posttest and Follow-up in the Video/Text and Live Conditions and the Respective Interaction Terms*

Dependent variable	Posttest (t <sub>2</sub> )						Follow-up (t <sub>3</sub> )					
	Video/text		Live		Interaction		Video/text		Live		Interaction	
	w <sub>v/t</sub> (SE)	p	w <sub>1</sub> (SE)	p	Δw (SE)	p	w <sub>v/t</sub> (SE)	p	w <sub>1</sub> (SE)	p	Δw (SE)	p
Understanding reconstruction	-0.10 (0.09)	.256	0.10 (0.08)	.209	0.20 (0.12)	.087	<b>-0.32 (0.10)</b>	<b>.001</b>	0.09 (0.12)	.429	<b>0.41 (0.15)</b>	<b>.006</b>
Understanding deconstruction	<b>-0.22 (0.06)</b>	<b>.000</b>	0.01 (0.05)	.904	<b>0.22 (0.08)</b>	<b>.004</b>	-0.13 (0.08)	.079	-0.02 (0.06)	.794	0.12 (0.10)	.240
Understanding oral history	-0.16 (0.07)	.027	0.12 (0.09)	.169	<b>0.29 (0.11)</b>	<b>.011</b>	-0.17 (0.10)	.082	<b>0.19 (0.08)</b>	<b>.014</b>	<b>0.35 (0.13)</b>	<b>.006</b>

*Note.* Posttest:  $N_{video/text} = 516$ ,  $N_{live} = 250$ ; Fit indices:  $\chi^2(195) = 316.72$ ,  $p < .001$ , RMSEA = 0.04, CFI = .96, SRMR = .05; Follow-up:  $N_{video/text} = 513$ ,  $N_{live} = 247$ ; Fit indices:  $\chi^2(195) = 317.86$ ,  $p < .001$ , RMSEA = 0.04, CFI = 0.96, SRMR = 0.05; w<sub>v/t</sub>, w<sub>1</sub> = group-specific regression coefficient; Δw = w<sub>1</sub> – w<sub>v/t</sub>, estimate of interaction term; SE = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold.

**Figure 4**

*Differences in the Treatment Effect for Students with Low and High Comfort in Social Situations With Regard to the Respective Competence Measures at Posttest and Follow-up*



### Predicting Behavioral Engagement from Students' Comfort in Social Situations

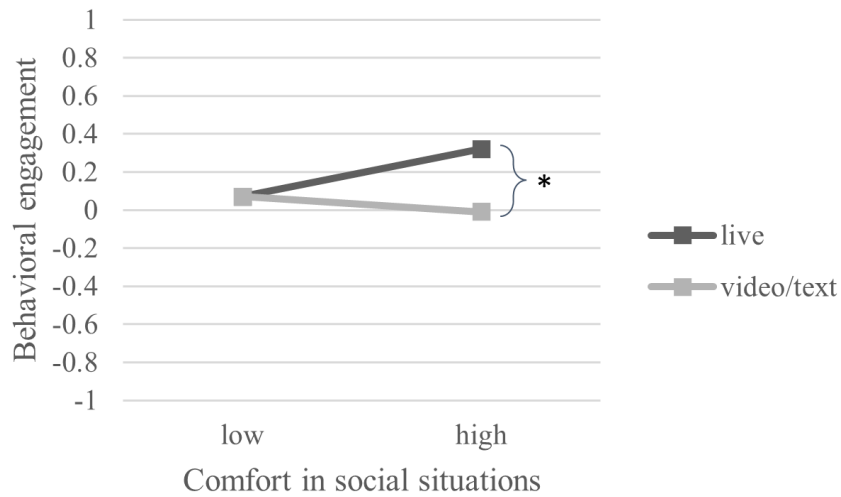
Our set of regression analyses described so far supported ATIs between students' comfort in social situations and the presence of eyewitnesses in history lessons. Following our second research question, we were interested in the effect of students' comfort in social situations on behavioral engagement as a process variable. To investigate these effects, we expanded the first set of models into mediation models with behavioral engagement during the teaching unit as a mediator (see Figure 2). The expanded multigroup models had good model fit indices: posttest:  $\chi^2(325) = 480.84, p < .001, RMSEA = 0.04, CFI = 0.96; SRMR = 0.05;$  follow-up:  $\chi^2(325) = 487.50, p < .001, RMSEA = 0.04, CFI = 0.96, SRMR = 0.05.$

We examined the group-specific effect of students' dispositional comfort in social situations on behavioral engagement ( $a_{v/t}, a_1$ ), and the difference between groups for our second research question. The effect of students' comfort on engagement in the video/text condition was close to zero ( $a_{v/t} = -0.01, SE = 0.04, p = .879$ ), while the effect in the live condition was small ( $a_1 = 0.13, SE = 0.06, p = .047$ ). Accordingly, the difference between the two groups in the effect of comfort on engagement was small (Figure 2:  $a_1 - a_{v/t}: \Delta a = 0.13, SE = 0.08, p = .091$ ). The results indicated that the higher students' comfort in social situations, the higher their behavioral engagement in the teaching unit with the live eyewitness and that students' comfort did not matter in the video/text condition with regard to their behavioral engagement. However, after controlling for the FDR, none of these findings were statistically significant.

To facilitate our interpretations, we estimated the group differences in behavioral engagement for students with low ( $-1 SD$ ) and high ( $+1 SD$ ) comfort. The conditional effects plot (Figure 5) illustrates that behavioral engagement during the whole teaching unit differed in particular for students with high comfort in social situations. Figure 5 indicates that the video/text condition did not seem to have a differential impact on engagement for either student group (as shown by the rather small difference in behavioral engagement between students with low/high comfort in social situations), whereas the live condition was particularly good for fostering behavioral engagement for students with high comfort in social situations.

**Figure 5**

*Behavioral Engagement for Students With Low (-1 SD) and High (+1 SD) Comfort in Social Situations in the Two Conditions*

**Behavioral Engagement as a Mediator of the ATI**

We began the examination of potential mediation effects, our last research question (RQ 3), by probing for the overall effect of behavioral engagement on the outcome measures without grouping; this allowed us to gauge whether behavioral engagement would be associated with the outcomes at all. The results indicated that behavioral engagement predicted the three competence measures from  $b = 0.19$  (understanding deconstruction) to  $b = 0.61$  (understanding reconstruction) at posttest and from  $b = 0.25$  (understanding deconstruction) to  $b = 0.60$  (understanding reconstruction) at follow up (Table 4). After controlling for the FDR, the association between behavioral engagement and understanding deconstruction at posttest turned out to be not statistically significant.

**Table 4**

*Regression Coefficients for Students' Behavioral Engagement in Predicting the Competence Measures at Posttest and Follow-Up Without Grouping*

Dependent variable	Posttest (t <sub>2</sub> )		Follow-up (t <sub>3</sub> )	
	b (SE)	p	b (SE)	p
Understanding reconstruction	<b>0.61 (0.09)</b>	<b>.000</b>	<b>0.60 (0.12)</b>	<b>.000</b>
Understanding deconstruction	0.19 (0.10)	.050	<b>0.25 (0.07)</b>	<b>.001</b>
Understanding oral history	<b>0.36 (0.11)</b>	<b>.001</b>	<b>0.44 (0.12)</b>	<b>.000</b>

*Note.*  $N = 769$ ; Models were controlled for all pretest measures as covariates; Fit indices:  $\chi^2(29) = 49.93, p = .009$ , RMSEA = 0.03, CFI = .99, SRMR = .02; b = estimate; SE = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold.

Finally, we took a look at the group-specific direct effects of comfort in social situations on outcomes, the indirect effects of comfort on outcomes via engagement, and the respective interaction terms (Table 5 and Table 6). The analyses revealed stronger—and negative—direct effects of students' comfort on outcome measures in the video/text condition than in the live condition at posttest and follow-up. After controlling for the FDR we found the effect of engagement on understanding deconstruction in the video/text condition at posttest ( $c_{v/t} = -0.21$ ,  $SE = 0.06$ ,  $p = .001$ ) and the effect of engagement on understanding reconstruction in the video/text condition at follow-up ( $c_{v/t} = -0.31$ ,  $SE = 0.10$ ,  $p = .002$ ) to be statistically significant. Results indicated further that there were no indirect effects of comfort in social situations on outcome measures via engagement in the video/text group at posttest and follow-up, but small indirect effects in the live condition at posttest and follow-up. Accordingly, we found small and negative first-stage moderation effects (i.e., differences in indirect effects between groups). These results were not statistically significant after controlling for the FDR. With regard to our last research question, we did not find support for moderated-mediation effects within the current study.<sup>1</sup>

<sup>1</sup> See Appendix D for additional analyses with factual knowledge and students' evaluations of their learning success and interest as dependent variables.



**Table 5**

*Group-Specific Direct Effects of Comfort in Social Situations on Outcome Measures ( $c_{v/t}$ ,  $c_l$ ) at Posttest and Follow-Up*

Dependent Variable	Posttest ( $t_2$ )				Follow-up ( $t_3$ )			
	Video/text		Live		Video/text		Live	
	$c_{v/t}$ ( <i>SE</i> )	<i>p</i>	$c_l$ ( <i>SE</i> )	<i>p</i>	$c_{v/t}$ ( <i>SE</i> )	<i>p</i>	$c_l$ ( <i>SE</i> )	<i>p</i>
Understanding reconstruction	-0.10 (0.10)	.294	0.03 (0.07)	.629	<b>-0.31</b> <b>(0.10)</b>	<b>.002</b>	0.03 (0.09)	.759
Understanding deconstruction	<b>-0.21</b> <b>(0.06)</b>	<b>.001</b>	-0.01 (0.05)	.861	-0.13 (0.08)	.094	-0.04 (0.06)	.480
Understanding oral history	-0.16 (0.07)	.024	0.08 (0.10)	.424	-0.16 (0.10)	.103	0.15 (0.08)	.050

Posttest:  $N_{video/text} = 516$ ,  $N_{live} = 250$ ; Fit indices:  $\chi^2(325) = 480.84$ ,  $p < .001$ , RMSEA = 0.04, CFI = 0.96; SRMR = 0.05; Follow-up:  $N_{video/text} = 513$ ,  $N_{live} = 247$ ; Fit indices:  $\chi^2(325) = 487.50$ ,  $p < .001$ ; RMSEA = 0.04; CFI = 0.96; SRMR = 0.05;  $c_{v/t}$ ,  $c_l$  = group specific regression coefficient; *SE* = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold.

**Table 6**

*Moderated Mediation: Group-Specific Indirect Effects of Comfort in Social Situations on Outcome Measures via Engagement ( $ab_{v/t}$ ;  $ab_l$ ) and the Respective Interaction Term ( $ab_l - ab_{v/t}$ ) for Outcome Measures at Posttest and Follow-Up*

Dependent Variable	Posttest (t <sub>2</sub> )						Follow-up (t <sub>3</sub> )					
	Video/text		Live		Interaction		Video/text		Live		Interaction	
	$ab_{v/t}$ (SE)	<i>p</i>	$ab_l$ (SE)	<i>p</i>	$\Delta ab$ (SE)	<i>p</i>	$ab_{v/t}$ (SE)	<i>p</i>	$ab_l$ (SE)	<i>p</i>	$\Delta ab$ (SE)	<i>p</i>
<b>Understanding reconstruction</b>	0.00 (0.03)	.879	0.08 (0.05)	.081	0.08 (0.05)	.123	-0.01 (0.03)	.751	0.08 (0.04)	.063	0.09 (0.05)	.091
<b>Understanding deconstruction</b>	0.00 (0.01)	.879	0.03 (0.02)	.102	0.03 (0.02)	.148	0.00 (0.01)	.750	0.04 (0.02)	.077	0.04 (0.02)	.100
<b>Understanding oral history</b>	0.00 (0.02)	.879	0.05 (0.03)	.094	0.05 (0.03)	.146	-0.01 (0.02)	.752	0.06 (0.03)	.087	0.06 (0.04)	.116

*Note.* Effects of engagement on outcome measures (path b in Figure 2) were set equal across groups (because they were not statistically significantly different). Posttest: Engagement on understanding reconstruction:  $b = 0.63$ ,  $SE = 0.09$ ,  $p < .001$ ; Engagement on understanding deconstruction:  $b = 0.24$ ,  $SE = 0.11$ ,  $p = .023$ ; Engagement on understanding oral history:  $b = 0.38$ ,  $SE = 0.11$ ,  $p = .001$ ; Follow-up: Engagement on understanding reconstruction:  $b = 0.21$ ,  $SE = 0.05$ ,  $p < .001$ ; Engagement on understanding deconstruction:  $b = 0.14$ ,  $SE = 0.04$ ,  $p < .001$ ; Engagement on understanding oral history:  $b = 0.18$ ,  $SE = 0.04$ ,  $p < .001$ ;  $ab_{v/t}$ ,  $ab_l$  = group-specific indirect effects coefficient;  $\Delta ab = ab_l - ab_{v/t}$  estimate of interaction term;  $SE$  = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold.

## Discussion

In this study, we investigated the role that students' dispositional comfort in social situations played in the effectiveness of a socially demanding instructional setting (i.e., a history teaching unit with a live eyewitness of the past). More concretely, we examined whether students' comfort in social situations moderated the effect of different instructional settings on students' historical thinking competencies and whether students' behavioral engagement in this teaching unit played a mediating role. The results indicated that students' comfort played a role when it came to the evaluation of a socially demanding treatment: Students with low comfort in social situations learned more in the video/text condition (i.e., in lessons that were taught "as usual"), whereas students with high comfort learned in both conditions comparably. The present investigation therefore yielded support for the aptitude-treatment paradigm. With regard to behavioral engagement during the teaching unit, the live treatment particularly fostered the engagement of students with high comfort: High comfort students engaged more in the live condition than in the video/text condition, whereas students with low comfort engaged in both conditions comparably. Hence, we found hints for the role that students' comfort played with regard to the process variable of engagement. We did not find empirical support for mediation or moderated mediation effects.

### **Students with Low and High Comfort in Social Situations Benefit from Different Settings**

In this study, we found hints for the differential effectiveness of a socially demanding instructional setting for different student groups. The results showed that whereas students with low dispositional comfort in social situations profited more from the video/text condition, students with high dispositional comfort either benefitted more from the live condition or benefitted from the two conditions comparably. The following section will offer potential explanations for the results of the current study.

In line with research on the meshing hypothesis, the assumption can be made that different levels of comfort in social situations are connected to different learning preferences and therefore have a differential impact on the growth of competence in different instructional settings (Fraser & Fisher, 1983). In particular, students with higher comfort in social situations can be expected to prefer to learn in a social setting, whereas students with lower comfort in social situations prefer to learn in lessons "as usual," working with text or video material rather than directly with a live person. Although psychological studies on the meshing hypothesis have not found strong support for the effects of learning styles on competencies (Pashler et al., 2009), these studies have focused primarily on verbal or visual learning preferences. To our

knowledge, preferences for learning in social situations have not been previously investigated. However, the results of this study can be interpreted as indicating that a match between comfort in social situations (indicating a certain learning preference) and the instructional setting leads to greater growth in competence. Furthermore, the results seem to indicate that especially for students who are less comfortable in social situations, a match is important, whereas students who are more comfortable find it easier to adapt to instructional settings that deviate from their preferences.

The learning goal of the teaching unit was rather abstract: Students were not supposed to learn details about the content of the eyewitness account but were rather supposed to learn, in general, that each eyewitness has his or her own personal perspective on the past. The eyewitness encounter was not expected to foster students' factual knowledge about the GDR but was instead supposed to foster students' historical thinking competencies because oral history accounts most commonly address personal experiences and perceptions rather than facts and figures (nevertheless, we explored additional outcomes such as factual knowledge in Appendix D). As Bertram et al. (2017) argued, this teaching goal might be especially difficult with regard to the emotional learning material that an oral history interview often reflects. For example, Sabrow (2012) claimed that when the experience of the eyewitness is more devastating, the person's "aura of authenticity" becomes stronger, and therefore, it has a more overwhelming effect on the listener. Researchers from crime psychology have observed something similar: Landström et al. (2005) compared evaluations of live witnesses in court with evaluations of equivalent videos by mock jurors and found, in line with prior studies, that live witnesses were evaluated with more leniency and were found to be more eloquent and positive than their video equivalent, independent of the actual truths or lies the witness was telling. The authors explained this by the participants' physical proximity to the live person, which led the observer to identify more with the witness and to provide a more positive evaluation. Taking on the role of a critical judge is easier without this proximity (Landström et al., 2005). Our findings suggest that higher comfort in social situations enables students to maintain a critical distance when reflecting on the eyewitness account regardless of physical proximity. The results confirm Pfundmair et al.'s (2017) findings: Students who show self-conscious behavior in social situations are less easily persuaded by the story, whereas students with lower comfort in social situations are more easily persuaded, especially in situations with a credible source (Pfundmair et al., 2017). By contrast, students with lower comfort seem to lack the ability to keep their distance as a critical judge, identify more with the person, and accept the story as trustworthy if it is displayed in the proximal live setting. Supporting this assumption, Cohen

(2001) suggested that identifying with a story character is easier to do in a noninteractive setting because the listener is more easily absorbed into the story and is not required to think about questions on a meta-level. Students with low comfort in social situations can be expected to avoid interacting with the live eyewitness and rather perceive the vivid and emotional story about the past without thinking about questions and therefore find it more challenging to scrutinize the story.

To probe for further explanations for the moderation effects, we investigated behavioral engagement during the teaching unit as a possible mediator of these effects. Although behavioral engagement played a role in predicting the outcome measures, and students with high comfort in social situations were found to be more engaged in the live condition than in the video/text condition, we did not find that behavioral engagement was a mediator of comfort level on competencies in either of the treatment conditions. The statistical power of the mediation effects would potentially be larger if the operationalization of the process variable behavioral engagement were specific to the respective lesson and not generic for the whole teaching unit. The association of students' dispositional comfort in social situations with behavioral engagement as a process measure emphasizes the need to investigate processes within the ATI paradigm.

### **Implications for Research on ATIs**

In their study, Bertram et al. (2017) analyzed the main effects of the intervention on students' historical thinking competencies and found that students in the video/text condition learned more with regard to understanding deconstruction and understanding oral history at posttest in comparison with students in the live condition. The current study expanded on these results by applying differential analyses and showed that the effects differed for students with different characteristics, that is, for those with different levels of dispositional comfort in social situations. These findings emphasize the importance of taking into account students' heterogeneity when examining the effectiveness of an instructional treatment rather than looking at only main effects.

In line with Snow (1992), this study highlights the importance of addressing not only cognitive moderators but also moderators that can be expected to be important in the intervention of interest. The readiness of a person to deal with a certain situation is determined not only by his or her cognitive ability but probably also by the person's conative or affective characteristics. The results of this study should encourage researchers to conduct ATI research

with an expanded focus and to take into account moderators that can be expected to be important in the intervention.

Furthermore, studies following the ATI paradigm often lack explanatory power (e.g., Ackerman, 2003). By including behavioral engagement as a mediating variable, we strengthened our analytical models, and we used a sample size (a total of 756 participants) that profoundly exceeded the sample sizes that have been used in a large number of ATI studies. In fact, we found effects of the moderator on engagement in the live condition, which stresses the importance of looking at process variables. At the same time, however, even in our study with a fairly large number of participants, we did not find moderated mediation effects. The results leave room for the exploration of additional mediators that can explain the differential effectiveness of the treatment for students who differ in their levels of comfort in social situations. Keeping the possible explanations of our findings in mind, a feeling of proximity might be able to explain the effects. However, the differential treatment effects for students with different levels of comfort in social situations need further investigation.

### **Limitations and Directions for Further Research**

Several research endeavors have recently emphasized the importance of looking at the heterogeneity of treatment effects in educational interventions (e.g., Reardon & Stuart, 2017). Our study contributes to research on the heterogeneity of treatment effects because we investigated affective moderator variables and included behavioral engagement as a mediating variable to further the understanding of the differential dynamics of an instructional treatment. However, as with all empirical studies, some limitations should be kept in mind.

First, it is important to understand the nature of our outcome measures. In our study, profiting from the treatment meant gaining insights into core constructs of historical thinking. The goal of the teaching unit was to foster historical thinking competencies. The three outcome measures used to assess aspects of students' historical thinking competencies, in particular epistemological principles of history, required different levels of abstraction: understanding reconstruction and understanding deconstruction refer to a rather high level of abstraction, whereas understanding oral history is more easily connected to the immediate eyewitness encounter. Although these measures offer differentiated insights, their different focuses complicate the ability to draw a conclusion about a general finding. In addition, insights into epistemological principles of history address only one aspect (albeit a central one; see Körber & Meyer-Hamme, 2015) of historical thinking competencies. Learning with oral history interviews might also be useful for fostering different dimensions of historical thinking

competencies that were not investigated in our study—such as competencies in orientation or methodological competencies (Körber & Meyer-Hamme, 2015) or the perceived authenticity of the source, teaching preferences, or social competencies such as empathy or perspective taking. However, by investigating the effects of the teaching unit on process variables such as behavioral engagement, we were able to show that there are other outcomes that might possibly be important.

A second limitation of our study involves the mediator variable behavioral engagement. To measure this variable, students were asked about their engagement during the teaching unit as a whole rather than about their engagement in the specific lesson with the eyewitness. This might explain the lack of findings in the mediation analyses. Additionally, the mediator was assessed at the same time point as the dependent variable for the posttest analyses. Finally, we relied on students' self-reports for the assessment of behavioral engagement. Although modern technologies increasingly allow for the use of additional assessments of behavioral engagement, self-reports of behavioral engagement are still used in many studies (e.g., Fredricks et al., 2004) because they are fairly easy to assess, have high acceptance among students, parents, and teachers, and have been shown to predict several important outcome variables (e.g., Reyes et al., 2012; Wang et al., 2018). Hence, at the current stage of research on ATI effects, we believe it is justifiable to use self-reports of behavioral engagement. We chose to use behavioral engagement in this form because we expected the students to be influenced by the information in the eyewitness format at the beginning of the teaching unit.

Finally, as is also true for many other studies on ATI effects, the current study was not initially designed to probe for ATI effects. Consequently, the sample sizes were originally calculated to detect main effects of the intervention. However, the design of the study and the sample size still provided a very good opportunity for running the analyses used in our study as shown by the post hoc power analyses. More specifically, because of the rather large number of students in the sample, there was adequate power for detecting moderation effects, whereas one might argue that a higher level of power was needed for the moderated mediation analysis. On a more general level, we believe that attempts to replicate our findings on the important role of comfort in learning situations are warranted to support the generalizability of our findings. To this end, researchers may check existing data sets for adequate variables or plan new studies.

**Conclusion**

In summary, the ATI analyses showed that especially students with lower dispositional comfort in social situations can be expected to profit more from lessons as usual with regard to their historical thinking competencies. By contrast, behavioral engagement did not mediate the effect of the Comfort  $\times$  Treatment interaction on historical thinking competencies. However, we found students with high comfort in social situations to engage more in the live condition than in the video/text condition in comparison with their classmates with lower comfort in social situations.



---

**References**

- Ackerman, P. L. (2003). Aptitude complexes and trait complexes. *Educational Psychologist*, 38(2), 85–93.
- Andrich, D. (1982). An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Educational Research and Perspectives*, 9(1), 95–104.
- Angvik, M., & von Borries, B. (Eds.) (1997). *Youth and history. A comparative European survey on historical consciousness and political attitudes among adolescents*. Körber-Stiftung.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
- Bertram, C., Wagner, W., & Trautwein, U. (2017). Learning historical thinking with oral history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54(3), 444–484. <https://doi.org/10.3102/0002831217694833>
- Bloom, H. S., & Spybrook, J. (2017). Assessing the precision of multisite trials for estimating the parameters of a cross-site population distribution of program effects. *Journal of Research on Educational Effectiveness*, 10(4), 877–902. <https://doi.org/10.1080/19345747.2016.1271069>
- Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 372–398. <https://doi.org/10.1080/10705511.2012.687671>
- Busby, R. S., & Hubbard, J. D. (2007). Using local oral history in the elementary classroom. *Social Studies Research and Practice*, 2(3), 367–386.
- Caspi, A., Chajut, E., Saporta, K., & Beyth-Marom, R. (2006). The influence of personality on social participation in learning environments. *Learning and Individual Differences*, 16, 129–144. <https://doi.org/10.1016/j.lindif.2005.07.003>
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12(3), 493–505. <https://doi.org/10.1207/s15328007sem1203>
- Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication and Society*, 4(3), 245–264. [https://doi.org/10.1207/S15327825MCS0403\\_01](https://doi.org/10.1207/S15327825MCS0403_01)

- Connolly, J. (1989). Social self-efficacy in adolescence: Relations with self-concept, social adjustment, and mental health. *Canadian Journal of Behavioural Science*, 21(3), 258–269. <https://doi.org/10.1037/h0079809>
- Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, 43(3), 161–173. <https://doi.org/10.1080/00461520802178466>
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12(11), 671–684. <https://doi.org/10.1037/h0043943>
- Cronbach, L. J., & Webb, N. (1975). Between-class and within-class effects in a reported aptitude  $\times$  treatment interaction: Reanalysis of a study by G. L. Anderson. *Journal of Educational Psychology*, 67(6), 717–724. <https://doi.org/10.1037/0022-0663.67.6.717>
- DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., & Bellinger, J. (2018). A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade. *Journal of Educational Psychology*, 110(1), 1–16. <https://doi.org/10.1037/edu0000191>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Dutt-Doner, K. M., Allen, S., & Campanaro, K. (2016). Understanding the impact of using oral histories in the classroom. *The Social Studies*, 107(6), 257–265. <https://doi.org/10.1080/00377996.2016.1221792>
- Edwards, J. R. (2009). Seven deadly myths of testing moderation in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends* (pp. 143–164). Routledge. <https://doi.org/10.4324/9780203867266>
- Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, 12(1), 1–22. <https://doi.org/10.1037/1082-989X.12.1.1>
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Educational Psychology Papers and Publications*, 8(3), 430–457. [https://doi.org/10.1207/S15328007SEM0803\\_5](https://doi.org/10.1207/S15328007SEM0803_5)
- Fend, H., & Prester, H.-G. (1986). *Bericht aus dem Projekt "Entwicklung im Jugendalter"* [Report from the project "development in adolescence"]. Konstanz.
- Flunger, B., Mayer, A. & Umbach, N. (2019). Beneficial for some or for everyone? Exploring

- the effects of an autonomy-supportive intervention in the real-life classroom. *Journal of Educational Psychology*, *111*(2), 210–234.  
<https://doi.org/http://dx.doi.org/10.1037/edu0000284>
- Fraser, B. J., & Fisher, D. L. (1983). Use of actual and preferred classroom environment scales in person-environment fit research. *Journal of Educational Psychology*, *75*(2), 303–313. <https://doi.org/10.1037/0022-0663.75.2.303>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109.  
<https://doi.org/10.3102/00346543074001059>
- Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2018). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology*, *110*(5), 628–639.  
<https://doi.org/http://dx.doi.org/10.1037/edu0000228>
- Fuchs, D., & Fuchs, L. S. (2019). On the importance of moderator analysis in intervention research: An introduction to the special issue. *Exceptional Children*, *85*(2), 1–3.  
<https://doi.org/10.1177/0014402918811924>
- Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., ... Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, *106*(2), 499–514.  
<https://doi.org/10.1037/a0034341>
- Furnham, A. (2012). Learning styles and approaches to learning. *APA Educational Psychology Handbook*, *2*, 59–81. <https://doi.org/10.1037/13274-003>
- Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, *108*(1), 82–97.  
<https://doi.org/10.1037/edu0000053>
- Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing Type S (sign) and Type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.  
<https://doi.org/10.1177/1745691614551642>
- Hattie, J. (2009). *Visible Learning—A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hedges, L. V., & Schauer, J. (2018). Randomised trials in education in the USA. *Educational Research*, Advance online publication. <https://doi.org/10.1080/00131881.2018.1493350>
- Hogervorst, S. (2018). Distanced by the screen. Student history teachers and video archives of Second World War interviews in the Netherlands. In W. Dreier, A. Laumer, & M. Wein

- (Eds.), *Explorations of good practice in educational work with video testimonies of victims of national socialism* (pp. 145–152). Stiftung “Erinnerung, Verantwortung und Zukunft.”
- Jerusalem, M., & Klein-Hessling, J. (2002). Soziale Kompetenz [social competence]. *Zeitschrift Für Psychologie / Journal of Psychology*, *210*(4), 164–174.  
<https://doi.org/10.1026//0044-3409.210.4.164>
- Jerusalem, M., & Klein-Hessling, J. (2009). Mental health promotion in schools by strengthening self-efficacy. *Health Education*, *109*(4), 329–341.  
<https://doi.org/10.1108/09654280910970901>
- Kagan, J. (1997). Temperament and the reactions to unfamiliarity. *Child Development*, *68*(1), 139–143.
- Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude-treatment interaction approach to writing-to-learn. *Learning and Instruction*, *18*(4), 379–390.  
<https://doi.org/10.1016/j.learninstruc.2007.07.004>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.
- Körber, A., & Meyer-Hamme, J. (2015). Historical thinking, competencies, and their measurement. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 89–101). Routledge.
- Korem, A. (2018). Supporting shy students in the classroom: A review. *European Psychologist*. <https://doi.org/10.1027/1016-9040/a000343>
- Krause, S., Back, M. D., Egloff, B., & Schmukle, S. C. (2016). Predicting self-confident behaviour with implicit and explicit self-esteem measures. *European Journal of Personality*, *30*, 648–662. <https://doi.org/10.1002/per.2076>
- Landström, S., Granhag, P. A., & Hartwig, M. (2005). Witnesses appearing live versus on video: Effects on observers’ perception, veracity assessments and memory. *Applied Cognitive Psychology*, *19*(7), 913–933. <https://doi.org/10.1002/acp.1131>
- Lanman, B. A., & Wendling, L. M. (2006). *Preparing the next generation of oral historians*. Altamira Press.
- Leung, K. C., Marsh, H. W., Yeung, A. S., & Abduljabbar, A. S. (2015). Validity of social, moral and emotional facets of Self-Description Questionnaire II. *Journal of Experimental Education*, *83*(1), 1–23. <https://doi.org/10.1080/00220973.2013.876229>
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, *31*(3), 313–327.

- <https://doi.org/10.1177/017084068800900203>
- Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly, 19*(2), 119–137. <https://doi.org/10.1080/10573560308223>
- Marzano, R., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works*. Association for Supervision and Curriculum Development.
- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114–140. <https://doi.org/https://doi.org/10.1037/met0000078>
- Medaille, A., & Usinger, J. (2019). Engaging quiet students in the college classroom. *College Teaching, 67*(2), 130–137. <https://doi.org/10.1080/87567555.2019.1579701>
- Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121–131. <https://doi.org/10.1037/a0033546>
- Muenks, K., Wigfield, A., Yang, J. S., & O’Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students’ personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology, 109*(5), 599–620. <https://doi.org/10.1037/edu0000153>
- Murphy, L., Eduljee, N. B., Croteau, K., & Parkman, S. (2017). Extraversion and introversion personality type and preferred teaching and classroom participation: A pilot study. *Journal of Psychosocial Research, 12*(2), 437–450.
- Muthén, L. K., & Muthén, B. O. (n.d.). *Mplus user’s guide* (8th ed.). Muthén & Muthén.
- Nguyen, T. D., Cannata, M., & Miller, J. (2018). Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *Journal of Educational Research, 111*(2), 163–174. <https://doi.org/10.1080/00220671.2016.1220359>
- Oral History Association. (2009). *Principles for oral history and best practices for oral history*. Retrieved from <http://www.oralhistory.org/about/principles-and-practices/>
- Paas, F., van Merriënboer, J. J. G., & van Gog, T. A. J. M. (2012). Designing instruction for the contemporary learning landscape. *APA Educational Psychology Handbook, 3*, 335–357. <https://doi.org/10.1037/13275-013>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., ... Allen, M. (2017). Teachers’ instructional adaptations: A research synthesis. *Review of Educational Research, 88*(2), 205–242. <https://doi.org/10.3102/0034654317743198>

- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles—concepts and evidence. *Psychological Science in the Public Interest*, *9*(3), 105–119.
- Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*(1), 83–98. <https://doi.org/10.1037/0022-0663.99.1.83>
- Pfundmair, M., Aydin, N., & Frey, D. (2017). Whatever? The effect of social exclusion on adopting persuasive messages. *Journal of Social Psychology*, *157*(2), 181–193. <https://doi.org/10.1080/00224545.2016.1192098>
- Reardon, S. F., & Stuart, E. A. (2017). Editors' introduction: Theme issue on variation in treatment effects. *Journal of Research on Educational Effectiveness*, *10*(4), 671–674. <https://doi.org/10.1080/19345747.2017.1386037>
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, *104*(3), 700–712. <https://doi.org/10.1037/a0027268>
- Rosenzweig, R., & Thelen, D. P. (1998). *The presence of the past: Popular uses of history in American life*. Columbia University Press.
- Ryu, E., & Cheong, J. (2017). Comparing indirect effects in different groups in single-group and multi-group structural equation models. *Frontiers in Psychology*, *8*, 1–14. <https://doi.org/10.3389/fpsyg.2017.00747>
- Sabrow, M. (2012). Der Zeitzeuge als Wanderer zwischen den Welten [The contemporary witness as wayfarer between two worlds]. In M. Sabrow & N. Frei (Eds.), *Die Geburt des Zeitzeugen nach 1945* (pp. 13–32). Wallstein Verlag.
- Snow, R. E. (1977). Individual differences and instructional theory. *Educational Researcher*, *6*(10), 11–15.
- Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, *27*(1), 5–32.
- Snow, R. E., Corno, L., & Cronbach, L. J. (2002). *Remaking the concept of aptitude—extending the legacy of Richard E. Snow*. Erlbaum.
- Snow, R. E., Kanfer, R., & Cudeck, R. (1989). Cognitive-conative aptitude interactions in learning. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences*. (pp. 435–475). Lawrence Erlbaum Associates Inc.
- Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, *76*(3), 347–376.

- <https://doi.org/10.1037/0022-0663.76.3.347>
- Sochatzy, F. (2010). Zeitzeugen im Geschichtsunterricht. In M. Ventzke, S. Mebus, W. Schreiber, B. von Borries, A. Körber, C. Pfluger, ... J. Polak (Eds.), *Geschichte denken statt pauken in der Sekundarstufe II. 20 Jahre nach der friedlichen Revolution: Deutsche und europäische Perspektiven im gymnasialen Geschichtsunterricht*. Sächsisches Bildungsinst.
- Spybrook, J., Kelcey, B., & Dong, N. (2016). Power for detecting treatment by moderator effects in two- and three-level cluster randomized trials. *Journal of Educational and Behavioral Statistics, 41*(6), 605–627. <https://doi.org/10.3102/1076998616655442>
- Stemler, S. E., & Sternberg, R. J. (2013). The assessment of aptitude. In K. F. Geisinger (Ed.), *APA Handbook of Testing and Assessment in Psychology* (Vol. 3. Testing, pp. 281–296). <https://doi.org/10.1037/14049-013>
- Tras, Z. (2016). Analysis of social self efficacy and emotional intelligence in university students. *Social Science and Education Research Review, 113*, 107–113.
- Trofanenko, B. (2017). “We tell stories”: Oral history as a pedagogical encounter. In K. R. Llewellyn & N. Ng-A-Fook (Eds.), *Oral history and education* (pp. 149–165). Palgrave Macmillan. <https://doi.org/10.2307/3119791>
- Trškan, D. (2016). *Oral history education—dialogue with the past*. Slovenian National Commission for UNESCO.
- Wang, M. Te, Kiuru, N., Degol, J. L., & Salmela-Aro, K. (2018). Friends, academic achievement, and school engagement during adolescence: A social network approach to peer influence and selection effects. *Learning and Instruction, 58*(July 2016), 148–160. <https://doi.org/10.1016/j.learninstruc.2018.06.003>
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond “ $p < 0.05$ .” *American Statistician, 73*(sup1), 1–19. <https://doi.org/10.1080/00031305.2019.1583913>
- Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: Questioning kappa-squared as mediation effect size measure. *Psychological Methods, 20*(2), 193–203. <https://doi.org/10.1037/met0000040>
- Whitman, G. (2004). *Dialogue with the past*. Altamira Press.
- Yeh, Y.-C. (2012). Aptitude-treatment interaction. In N. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 295–298). Springer. <https://doi.org/10.1007/978-1-4419-1428-6>
- Yow, V. (1997). “Do I like them too much?”: Effects of the oral history interview on the

interviewer and vice-versa. *The Oral History Review*, 24(1), 55–79.

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (30th ed.). University of California.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29(3), 663–676.

<https://doi.org/10.3102/00028312029003663>



---

### Appendix A

Social self-concept was assessed with four items (Fend & Prester, 1986; Leung et al., 2015).

1. When others hang out together in breaks, I am often not noticed. (r)
2. I am quite respected with my classmates.
3. I can do what I want, I can't get to my classmates somehow. (r)
4. Sometimes I feel a little bit like an outsider in my class. (r)

Social self-efficacy was assessed with eight items (Jerusalem & Klein-Hessling, 2002).

1. I can easily chat with fellow students, even if I don't know them.
2. When I want to get to know other people, I succeed, even if the others are rather reserved.
3. When I am with other people, I know what I have to do to be accepted.
4. Even in difficult conflicts with classmates, I can find a solution.
5. I manage to cope well even with difficult classmates.
6. If I have a quarrel with my friends, I can find a solution.
7. I am able to say what I think in class, even if everyone else does not agree with me.
8. I succeed well in convincing other classmates of my opinion.

Self-esteem was assessed with three items (Fend & Prester, 1986).

1. Overall, I have a positive image of myself.
2. Overall, I like myself.
3. All in all I have a very negative image of myself. (r)

### Appendix B

Before we estimated the regression models, we tested for measurement invariance for the latent variables in two steps (Chen et al., 2005). The goodness-of-fit statistics and statistics from the difference tests are displayed in Table 1: First, we tested for measurement invariance (configural, weak, and strong factorial invariance) in the first-order factors and compared the nested models with log-likelihood difference tests adjusted by the scaling correction factors (Bryant & Satorra, 2012). The results indicated that strong measurement invariance across groups for the first-order factors could be assumed, and hence, the loadings and intercepts could be set equal across groups (weak vs. strong invariance),  $p(\chi^2) = .584$ . The model reflecting strong measurement invariance fit the data well,  $\chi^2(79) = 107.98$ ,  $p = .017$ , RMSEA = 0.03, CFI = 0.98, SRMR = 0.04. Accordingly, in a second step, we calculated stepwise models to test for measurement invariance in the second-order factor across groups while at the same time assuming strong invariance in the measurement models of the first-order factors. The log-likelihood difference test implied that strong factorial invariance could be assumed, that is, the loadings and the intercepts of the first-order factors could be held constant across groups (e.g., weak vs. strong invariance),  $p(\chi^2) = .662$ . The model reflecting strong measurement invariance showed a good model fit,  $\chi^2(79) = 108.97$ ,  $p = .014$ , RMSEA = 0.03, CFI = 0.98, SRMR = 0.04.

**Table B1**

*Goodness-of-Fit Statistics for Tests of Measurement Invariance for First- and Second-Order Factor Models of Comfort in Social Situations*

	Constraints: first-order factors	Constraints: second-order factor	RMSEA	CFI	SRMR	Log-likelihood value	Number of parameters	<i>s</i>	$p(\chi^2)$	
<b>First-order factor models</b>										
1.1	Configural invariance	Factor structure	-	0.04	0.98	0.03	-5369.37	68	1.35	
1.2	Weak invariance	+ Factor loadings	-	0.04	0.98	0.04	-5371.44	61	1.36	.871
1.3	Strong invariance	+ Means/intercepts	-	0.03	0.98	0.04	-5376.48	51	1.39	.584
<b>Second-order factor models</b>										
2.1	Configural invariance	All of the above	Factor structure	0.04	0.98	0.04	-5374.49	56	1.39	
2.2	Weak invariance	All of the above	+ Factor loadings	0.03	0.98	0.04	-5375.81	54	1.39	.395
2.3	Strong invariance	All of the above	+ Means/intercepts	0.03	0.98	0.04	-5376.87	51	1.39	.662

*Note.*  $N_{video/text} = 481$ ,  $N_{live} = 230$ . The constraint columns display equality constraints across groups. Model comparison was calculated with a log-likelihood difference test adjusted by the scaling correction factors,  $s$  = scaling correction factor.

### Appendix C

To check whether the second-order factor comfort in social situations would lead to results that were similar to those from the first-order constructs, we calculated models for each first-order factor separately. All models showed a good model fit: Posttest: Social self-concept:  $\chi^2(53) = 63.70, p = .149$ , RMSEA = 0.02, CFI = 0.99, SRMR = 0.03; Social self-efficacy:  $\chi^2(37) = 53.90, p = .036$ , RMSEA = 0.04, CFI = 0.99, SRMR = 0.04; Self-esteem:  $\chi^2(37) = 47.28, p = .120$ , RMSEA = 0.03, CFI = 0.99, SRMR = 0.03; Follow-up: Social self-concept:  $\chi^2(53) = 89.69, p = .001$ , RMSEA = 0.04, CFI = 0.98, SRMR = 0.04; Social self-efficacy:  $\chi^2(37) = 46.22, p = .142$ , RMSEA = 0.03, CFI = 0.99, SRMR = 0.04; Self-esteem:  $\chi^2(37) = 49.26, p = .087$ , RMSEA = 0.03, CFI = .99, SRMR = .04. Table C1 shows detailed results of the multigroup analyses for all comfort-in-social-situation factors separately for posttest and follow-up. To aid interpretation, we provide Tables C2, C3, and C4, which show the pattern of results for the group-specific regression coefficients of the comfort-in-social-situation constructs for each outcome measure separately. The grey scale indicates the sizes of the effects as explained in the right-hand column and helps to capture the result pattern independent from the significance levels. The coefficients can be interpreted as effect sizes. Consistently, the pattern of results showed negative effects of students' self-evaluations on the competence measures in the video/text condition (Table C2) and positive effects in the live condition (Table C3). All interaction effects were negative (Table C4; live – video/text), 16 out of 18 results showed effects below Est. = .14. These detailed analyses confirmed the result pattern we also found by modeling comfort in social situations as a second-order factor.

**Table C1**

*Regression Coefficients for the Self-Evaluation Constructs in Predicting the Outcomes at Posttest and Follow-Up in the Video/Text and Live Conditions and their Interaction*

	Posttest (t <sub>1</sub> )						Follow-up (t <sub>2</sub> )					
	Video/text		Live		Interaction		Video/text		Live		Interaction	
	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>
<b>Understanding reconstruction</b>												
Social self-concept	-0.06 (0.08)	.433	0.06 (0.09)	.514	0.12 (0.06)	.319	<b>-0.37 (0.12)</b>	<b>.002</b>	0.09 (0.12)	.425	<b>0.46 (0.14)</b>	<b>.001</b>
Social self-efficacy	-0.10 (0.07)	.145	0.11 (0.08)	.199	0.21 (0.10)	.043	<b>-0.23 (0.09)</b>	<b>.009</b>	0.24 (0.11)	.037	<b>0.47 (0.15)</b>	<b>.002</b>
Self-esteem	-0.02 (0.06)	.735	<b>0.18 (0.06)</b>	<b>.003</b>	<b>0.20 (0.08)</b>	<b>.014</b>	<b>-0.16 (0.06)</b>	<b>.008</b>	0.11 (0.11)	.310	0.27 (0.13)	.028
<b>Understanding deconstruction</b>												
Social self-concept	<b>-0.20 (0.05)</b>	<b>.000</b>	-0.02 (0.06)	.712	<b>0.17 (0.06)</b>	<b>.006</b>	<b>-0.22 (0.08)</b>	<b>.005</b>	-0.01 (0.09)	.921	0.21 (0.11)	.068
Social self-efficacy	<b>-0.17 (0.06)</b>	<b>.002</b>	0.11 (0.06)	.087	<b>0.28 (0.09)</b>	<b>.001</b>	-0.08 (0.05)	.110	0.07 (0.08)	.382	0.15 (0.11)	.141
Self-esteem	-0.05 (0.05)	.264	-0.01 (0.06)	.860	0.04 (0.06)	.512	-0.04 (0.06)	.536	-0.01 (0.07)	.919	0.33 (0.10)	.731
<b>Understanding oral history</b>												
Social self-concept	<b>-0.16 (0.06)</b>	<b>.009</b>	0.13 (0.09)	.157	<b>0.29 (0.10)</b>	<b>.004</b>	-0.16 (0.08)	.045	<b>0.27 (0.07)</b>	<b>.000</b>	<b>0.43 (0.11)</b>	<b>.000</b>
Social self-efficacy	-0.09 (0.06)	.125	0.11 (0.16)	.485	0.21 (0.16)	.192	-0.16 (0.08)	.045	0.22 (0.11)	.040	<b>0.38 (0.15)</b>	<b>.009</b>
Self-esteem	-0.07 (0.06)	.256	<b>0.16 (0.06)</b>	<b>.009</b>	<b>0.30 (0.09)</b>	<b>.008</b>	-0.07 (0.08)	.370	0.17 (0.07)	.019	0.24 (0.11)	.027

*Note.* Posttest:  $N_{Video/text} = 513$ ,  $N_{Live} = 247$ ; Follow-up:  $N_{Video/text} = 516$ ,  $N_{Live} = 250$ ; Interaction term = live – video/text; Est. = estimate; SE = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold (54 tests).

**Table C2***Result Pattern for the Video/Text Condition*

Video/text	Posttest			Follow-up			-0.02 – -0.04
	Understanding reconstruction	Understanding deconstruction	Understanding oral history	Understanding reconstruction	Understanding deconstruction	Understanding oral history	
social self-concept	-0.06	-0.20	-0.16	-0.37	-0.22	-0.16	-0.05 – -0.14
social self-efficacy	-0.07	-0.15	-0.07	-0.23	-0.08	-0.16	-0.15 – -0.24
self-esteem	-0.02	-0.05	-0.07	-0.16	-0.04	-0.07	≤ -0.25

**Table C3***Result Pattern for the Live Condition*

Live	Posttest			Follow-up			≤ 0.04
	Understanding reconstruction	Understanding deconstruction	Understanding oral history	Understanding reconstruction	Understanding deconstruction	Understanding oral history	
social self-concept	0.06	0.02	0.13	0.09	-0.01	0.27	0.05 – 0.14
social self-efficacy	0.09	0.10	0.09	0.24	0.07	0.22	0.15 – 0.24
self-esteem	0.18	-0.01	0.16	0.11	-0.01	0.17	≤ 0.25

**Table C4***Result Pattern for the Interaction Terms (Live – Video/Text)*

Interaction	Posttest			Follow-up			≤ 0.04
	Understanding reconstruction	Understanding deconstruction	Understanding oral history	Understanding reconstruction	Understanding deconstruction	Understanding oral history	
social self-concept	0.12	0.17	0.29	0.46	0.21	0.43	0.05 – 0.14
social self-efficacy	0.21	0.28	0.21	0.47	0.15	0.38	0.15 – 0.24
self-esteem	0.20	0.04	0.30	0.27	0.33	0.24	≤ 0.25

### Appendix D

In addition to the three measures of historical thinking competencies, after the oral history intervention, Bertram et al. (2017) assessed students' factual knowledge about the German Democratic Republic as well as students' evaluations of their learning success and interest. Because facts about the GDR tended to be presented in the beginning of the teaching unit, we did not expect the moderation and mediation effects to differ between the treatment conditions. The encounter with the eyewitness account was intended to teach students about different perspectives in historical accounts and crucial constructs involved in historical thinking but not about "historical facts." With regard to students' reports on their learning success and interest, we expected students' comfort in social situations to be associated with their learning success and interest in the live condition more than in the video/text condition. Students who felt more comfortable when meeting the eyewitness in person could be expected to believe that they had learned more in the teaching unit.

Students' factual knowledge about the German Democratic Republic was assessed with fill-in-the-blank questions at all three measurement points. The test was divided into two parts at the pre- and posttest comprised of seven items each and combined in the follow-up test. A sample item is: "After opening the border between \_\_\_\_\_ and Austria, a mass exodus from the GDR began" (translated). Reliability ranged from Cronbach's  $\alpha = .46$  for Part A at pretest to Cronbach's  $\alpha = .83$  for the full test at follow-up.

Students' evaluations of their learning success and interest were assessed with three scales rated on a 4-point Likert scale (1 = *I don't agree at all* to 4 = *I completely agree*) at posttest and follow-up. Students rated their impression of how much they learned during the teaching unit on topic-related knowledge (content learning scale). A sample item from the five-item scale is: "In the oral history interview, I learned a lot about life in a dictatorial regime" (translated). Internal consistency of the scale was adequate (posttest: Cronbach's  $\alpha = .80$ ; follow-up: Cronbach's  $\alpha = .76$ ). Students evaluated how much the teaching unit enhanced their methodological learning about the essentials of history itself (thinking history scale). A sample item from the five-item scale is: "In the evaluation of the oral history interview, I learned about how historians work" (translated). Cronbach's  $\alpha$  was  $\alpha = .69$  at posttest and  $\alpha = .65$  at follow-up. Students reported on the teaching unit's motivation for themselves (interest in oral history scale). A sample item is: "Working with oral history interviews was exciting and interesting" (translated). This scale was also assessed with five items. Reliability was adequate (posttest: Cronbach's  $\alpha = .86$ ; follow-up: Cronbach's  $\alpha = .80$ ).

In line with the models in the current study, we calculated multiple-group models to investigate (a) students' comfort in social situations as a moderator of the effect of the treatment on factual knowledge as well as of students' self-evaluation of their learning success and interest at posttest and follow-up; (b) the moderated mediation by students' comfort in social situations of the treatment effect on factual knowledge and students' self-evaluations via behavioral engagement during the teaching unit. In the models with factual knowledge as the dependent variable, the pretest scores were used as covariates. Students' self-report scales were modeled jointly, and the pretest scores for the historical thinking competencies served as covariates. In line with the main results of the current study, we report group-specific regression coefficients and their interaction for the moderation analysis and the group-specific indirect effects as well as their interaction for the mediated moderation.

Results of the moderation analyses are displayed in Table B1. As expected, there were no effects of students' comfort in social situations on factual knowledge in any of the conditions. With regard to students' reports of learning success and interest, we found effects close to zero for all three measures in the video/text condition at posttest and follow-up. In the live condition, the effect sizes were larger than in the video/text condition with regard to students' comfort in social situations on students' report on content learning, methodological learning, and interest at posttest and follow-up. However, after controlling for FDR, none of the effects were statistically significant. We did not find support for a moderation by students' comfort in social situations of the treatment effect on students' report on learning success and interest.

Group-specific indirect effects of comfort in social situations on factual knowledge and students' evaluation of their learning success and interest via engagement and the respective moderated mediation are displayed in Table B2. Consistently, we found no indirect effects in the video/text condition or in the live condition. The effects of the moderated mediation ranged from Est. = 0.06 to Est. = 0.11 with regard to students' reports of their learning success and interest. None of these moderated mediation effects were statistically significant.



**Table D1**

*Moderation Analysis: Regression Coefficients for Comfort in Social Situations on the Outcomes at Posttest and Follow-Up in the Video/Text and Live Conditions and the Respective Interaction Terms*

Dependent variable	Posttest (t <sub>1</sub> )						Follow-up (t <sub>2</sub> )					
	Video/text		Live		Interaction		Video/text		Live		Interaction	
	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>
Factual knowledge	-0.02 (0.02)	.140	-0.03 (0.02)	.120	0.00 (0.02)	.918	-0.01 (0.01)	.308	0.01 (0.01)	.644	0.01 (0.01)	.315
<b>Student report</b>												
Content learning	0.01 (0.06)	.901	0.10 (0.04)	.012	0.09 (0.07)	.214	0.01 (0.04)	.731	0.13 (0.05)	.015	0.11 (0.07)	.095
Thinking history	0.01 (0.05)	.785	0.09 (0.04)	.047	0.07 (0.06)	.244	0.04 (0.04)	.257	0.10 (0.04)	.012	0.06 (0.06)	.334
Interest in oral history	0.02 (0.05)	.750	0.11 (0.04)	.006	0.09 (0.07)	.170	0.08 (0.04)	.025	0.17 (0.08)	.026	0.09 (0.09)	.289

*Note.* Factual knowledge posttest:  $N_{video/text} = 516$ ,  $N_{live} = 250$ ;  $\chi^2(116) = 179.66$ ,  $p < .001$ ; RMSEA = 0.04; CFI = 0.97; SRMR = 0.05; Follow-up:  $N_{video/text} = 513$ ,  $N_{live} = 247$ ;  $\chi^2(116) = 164.23$ ,  $p = .002$ ; RMSEA = 0.03; CFI = 0.98; SRMR = 0.05; Student report posttest:  $N_{video/text} = 508$ ,  $N_{live} = 249$ ; Fit indices student report:  $\chi^2(196) = 304.12$ ,  $p < .001$ ; RMSEA = 0.04; CFI = 0.97; SRMR = 0.05; Follow-up:  $N_{video/text} = 508$ ,  $N_{live} = 246$ ; Fit indices:  $\chi^2(196) = 323.29$ ,  $p < .001$ ; RMSEA = 0.04; CFI = 0.96; SRMR = 0.05; Interaction term = live – video/text; Est. = estimate; SE = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold (24 tests).

**Table D2**

*Moderated Mediation: Group-Specific Indirect Effects of Comfort in Social Situations on the Outcome Measures via Engagement and the Respective Interaction Term for Factual Knowledge and Students' Evaluation of Learning Success at Posttest and Follow-Up*

Dependent variable	Posttest (t <sub>1</sub> )						Follow-up (t <sub>2</sub> )					
	Video/text		Live		Interaction		Video/text		Live		Interaction	
	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>	Est. (SE)	<i>p</i>
Factual Knowledge	0.00 (0.00)	.956	0.01 (0.01)	.123	0.01 (0.01)	.182	0.00 (0.00)	.893	0.01 (0.01)	.229	0.01 (0.01)	.299
<b>Student report</b>												
Content learning	0.00 (0.02)	.970	0.05 (0.03)	.078	0.05 (0.04)	.151	0.00 (0.02)	.965	0.05 (0.03)	.056	0.05 (0.03)	.104
Thinking history	0.00 (0.02)	.970	0.07 (0.03)	.046	0.06 (0.04)	.118	0.00 (0.02)	.965	0.05 (0.03)	.048	0.05 (0.03)	.092
Interest in oral history	0.00 (0.03)	.970	0.07 (0.04)	.064	0.07 (0.04)	.132	0.00 (0.02)	.965	0.06 (0.03)	.039	0.06 (0.04)	.084

*Note.* Effects of engagement on outcome measures (path b in Figure 2) were tested by setting them equal between groups. Interaction term = live – video/text; Est. = estimate; SE = standard error; statistically significant coefficients controlling for the FDR at  $q^* = .05$  are printed in bold (24 tests).

---

# 7

## GESAMTDISKUSSION

---

## 7 Gesamtdiskussion

Das Lernen mit authentischen Geschichten im Geschichtsunterricht – mit Zeitzeug:innenberichten – stellt besondere Herausforderungen an das historische Denken: Durch den individuellen Zugang zur Vergangenheit sowie durch persönliche Erfahrungen über Zwänge und Herausforderungen in der vergangenen Zeit werden Lernende emotional angesprochen. Gleichzeitig schätzen die Lernenden die Geschichten als besonders glaubwürdig ein, insofern, als sie darauf vertrauen, dass diese die Wahrheit über die Vergangenheit erzählen können (Kapitel 1.2.1, Kapitel 1.2.3). Problematisch ist dabei, dass wir in der Gegenwart durch die Rekonstruktion vergangener Ereignisse niemals in der Lage sind, ein genaues Abbild der Vergangenheit herzustellen, sondern lediglich einzelne Narrative und Ausschnitte der Vergangenheit darstellen können, die nur in der Gesamtschau ein mehr oder weniger gesichertes Abbild der Vergangenheit erzeugen können. Diese Einsicht zu vermitteln ist ein Ziel des Geschichtsunterrichts (Kapitel 1.1.2). In dem Lernarrangement mit Zeitzeug:innenberichten wird deutlich, dass Emotionen den Zugang zur Geschichte verschließen können (Brauer & Lücke, 2013), wenn das Lernarrangement von der Lehrkraft nicht in eine Unterrichtseinheit, die eine quellenkritische Analyse instruiert, eingebettet wurde. (z. B. Rodenhäuser, 2013; Kapitel 1.2.3). Was aber führt genau zu diesem Konflikt zwischen der Wirkung des Lernarrangements und den eigentlichen Zielen des Geschichtsunterrichts? Was passiert zwischen der Begegnung mit den Zeitzeug:innenberichten und ihrer Wirkung?

Ich habe im theoretischen Rahmen herausgearbeitet, dass die mental ablaufenden Lernprozesse im Rahmen der Lernaktivität in vier Verarbeitungsmodi differenziert werden können, die sich in ein Koordinatensystem mit den Achsen *Persönliche Distanz zur Vergangenheit* und *Abstraktionsniveau* einordnen lassen (Kapitel 1.3.3). Für die Erklärung des Konflikts zwischen dem Ziel des Geschichtsunterrichts und der Wirkung des Lernarrangements gaben empirische Studien Hinweise darauf, dass insbesondere der distanzlos-involvierte Verarbeitungsmodus eine besondere Rolle spielt. Dieser zeichnet sich durch eine geringe wahrgenommene Distanz zur Vergangenheit und ein niedriges Abstraktionsniveau aus und schließt auch mit ein, dass Lernende durch die Begegnung mit den Zeitzeug:innenberichten überwältigt sind – folglich keine kritische Einordnung der Narrative erfolgt.

Die vorliegende Dissertation hat zum Ziel, das Lernarrangement mit Zeitzeug:innenberichten systematisch zu beleuchten, um Einblicke in die Blackbox der Lernprozesse, die durch das Lernarrangement initiiert werden, zu erhalten. Dabei sollten die Perspektiven von Geschichtsdidaktik und Empirischer Bildungsforschung miteinander verknüpft

werden. Denn geschichtsdidaktische Forschung nahm bisher nur selten das Lernarrangement mit empirischen Studien in den Blick, während Studien der Empirischen Bildungsforschung den Geschichtsunterricht nur selten zum Gegenstand ihrer Forschung machten. Der Zugang zu dem Lernarrangement über die zugrunde liegenden Lernprozesse wurde als besonders vielversprechend identifiziert, um die Wirkung des Lernarrangements zu verstehen. Das Angebots-Nutzungs-Modell zur Beschreibung unterrichtlicher Wirkungsweisen nach Helmke (2017) diene als generisches Rahmenmodell, das Anknüpfungspunkte für die Erforschung des Lernarrangements bot.

Im Folgenden werden zunächst die Ergebnisse der vier Dissertationsstudien kurz zusammengefasst (Kapitel 7.1) und mit Blick auf ihren Beitrag zur Geschichtsdidaktik (Kapitel 7.1.1) und zur Empirischen Bildungsforschung (Kapitel 7.1.2) diskutiert. Im Anschluss daran zeige ich Stärken und Grenzen der Dissertation auf (Kapitel 7.2), leite Implikationen für Forschung, schulische und außerschulische Bildungspraxis ab (Kapitel 7.3) und schließe mit einem allgemeinen Fazit (Kapitel 7.4).

## 7.1 Diskussion der Ergebnisse

Was machen die Lernprozesse in der Begegnung mit Zeitzeug:innenberichten aus? Diese Frage wurde in den Dissertationsstudien mit Blick auf (a) die Lernaktivität, (b) die Wahrnehmung und (c) das Zusammenspiel von Lernvoraussetzungen und Unterrichtsangebot untersucht. Es zeigte sich für den Bereich der Lernaktivität, dass die Verarbeitungsprozesse als distanzlos-involvierten Verarbeitungsmodus – dem „Aura-Erlebnis“ – verstanden werden können. Studie 1 entwarf ein konzeptionelles Modell, das die situationale Verarbeitung und das Involvement mit den Berichten anhand von kognitiven, affektiven und physischen Prozessen definierte. In Studie 2 überführte ich die Konzeptionalisierungen in ein Wirkmodell und setzte es in Beziehung zu dem Verarbeitungsmodus der kritisch-distanzierten Verarbeitung. In der Studie wurde als Grundlage für weitere empirische Forschungsarbeiten ein standardisiertes Messinstrument entwickelt, das den psychometrischen Gütekriterien der Objektivität, Reliabilität und Validität entspricht. Im Bereich der Wahrnehmung untersuchte Studie 3 in einem experimentellen Setting mit dem validierten Fragebogen aus Studie 2, ob die (wahrgenommene) Authentizität von Zeitzeug:innenberichten einen Einfluss auf die Verarbeitungsprozesse hatte. Es zeigte sich, dass das Framing der Videoclips als originaler Bericht, als nachgespielter Bericht oder als rekonstruierter Bericht insbesondere die affektiven Verarbeitungsprozesse beeinflusste. Lernende, die den originalen Bericht sahen, waren stärker emotional involviert als Lernende, die die anderen Berichte sahen. Schließlich untersuchte Studie 4 das Zusammenspiel von Lernvoraussetzungen und Unterrichtsangebot auf die Wirkungen, um einen weiteren Mechanismus der Lernprozesse in den Blick zu nehmen. Welchen Einfluss hat das Format der Berichte auf Kompetenzen historischen Denkens für Lernende mit unterschiedlichen Lernvoraussetzungen? Es zeigte sich, dass das soziale Wohlbefinden einen Einfluss auf die Wirkung des Formats der Zeitzeug:innenberichte auf historische Kompetenzen hatte. Studie 2 betrachtete den Zusammenhang von Merkmalen auf Personenebene (Geschlecht, Interaktion mit Zeitzeug:in) und auf Situationsebene (Intensität der Vorbereitung, Abstand zwischen Begegnung und Befragung) mit den Verarbeitungsprozessen, um ihre Rolle für die Lernprozesse zu ergründen. Hier zeigte sich, dass Lernende, die angaben, ein Mädchen zu sein, stärker involviert waren als Lernende, die angaben, ein Junge zu sein. Welchen Beitrag zur Forschung die Ergebnisse der vier Dissertationsstudien in den zwei beteiligten Disziplinen leisten, wird im Folgenden diskutiert.

### 7.1.1 Das Lernen mit Zeitzeug:innenberichten verstehen: Beitrag zur Geschichtsdidaktik

Das Angebots-Nutzungs-Modell für die Wirkungsweisen von Unterricht (Helmke, 2017) bietet vielfältige Anknüpfungspunkte, um das komplexe Unterrichtsgeschehen im Fachunterricht zu verstehen (Seidel, 2014). Die vorliegende Dissertation kann anhand des Angebots-Nutzungs-Modells sowie der Zugänge der Empirischen Bildungsforschung umfangreiche Einblicke in das Lernen mit Zeitzeug:innenberichten geben. Sowohl auf inhaltlicher als auch auf methodischer Ebene können die Dissertationsstudien Erkenntnisse für die Geschichtsdidaktik liefern. Konkret leistet die Dissertation einen Beitrag zu einem differenzierten Verständnis (a) der Lern- und Verarbeitungsprozesse, (b) der Rolle von Emotionen beim historischen Lernen sowie (c) der Authentizität von Zeitzeug:innenberichten. Schließlich leisten die Dissertationsstudien einen starken methodischen Beitrag zur geschichtsdidaktischen Forschung.

Im Zentrum der Dissertation standen die Lernprozesse beim Lernen mit Zeitzeug:innenberichten. Theoretische wie empirische Arbeiten in der Geschichtsdidaktik sowie Diskussionen aus der Bildungspraxis wiesen bereits häufig auf die Chancen und Risiken der Arbeit mit Zeitzeug:innen hin (z. B. Bertram, 2017; Henke-Bockschatz, 2007; Trofanenko, 2017; von Plato, 2001; Whitman, 2004; siehe auch Kapitel 1.2.3). Dabei konzentrierten sich die Arbeiten zu einem großen Teil auf die Wirkung der Begegnung mit den Zeitzeug:innenberichten. Die Lern- und Verarbeitungsprozesse wurden dabei nur in Ansätzen berücksichtigt (z. B. Brüning, 2018; Richardson, 2021). An diesen Punkt knüpfen drei der vier Dissertationsstudien an und legen den Fokus auf die Wahrnehmungs- und Verarbeitungsprozesse von Lernenden in diesem Lernarrangement. Dieser Betrachtung lag die Annahme zugrunde, dass die „Aura der Authentizität“ (Sabrow, 2012, S. 27) der Zeitzeug:innenberichte nicht ausschließlich von den Zeitzeug:innen selbst ausgeht, sondern durch eine Zuschreibung der Lernenden entsteht (Kapitel 1.3.2). Im Fokus der Betrachtungen stand das „Aura-Erlebnis“ – eine distanzlos-involvierte Verarbeitungsweise der Berichte – welches sich durch eine geringe wahrgenommene Distanz zur Vergangenheit und ein niedriges Abstraktionsniveau auszeichnet. In den Studien von Brüning (2018), Richardson (2021) sowie von Ballis und Schwendemann (2021) lassen sich Merkmale der Begegnung mit Zeitzeug:innen finden, wie beispielsweise eine gefühlte Verbindung zur Person und zur Vergangenheit (Brüning, 2018; Richardson, 2021), ein Eintauchen in die Geschichte (Brüning, 2018) und ein besonderes Vertrauen (Ballis & Schwendemann, 2021). Studie 1 und 2 differenzierten die Begegnung im Rahmen des „Aura-Erlebnisses“ noch weiter aus: So machten kognitive Prozesse, wie beispielsweise eine beson-

ders lebendige Vorstellung der damaligen Zeit, affektive Prozesse, wie beispielsweise das Bewegt Sein von den persönlichen Geschichten, sowie physische Prozesse, wie beispielsweise die sinnliche Wahrnehmung, der Geruch oder auch die Präsenz der Zeitzeug:innen, die Begegnung aus. Das widerspricht den bisherigen Studien nicht, sondern erweitert das Verständnis der Begegnung und gibt Anknüpfungspunkte, diese theoretisch zu diskutieren, zu erweitern und empirisch zu prüfen. Studie 2 konnte diese differenzierte Konzeptualisierung empirisch bestätigen und zudem Hinweise finden, dass diese distanzlos-involvierte Verarbeitung negativ mit der distanziert-analytischen Verarbeitung zusammenhängt. Dieses Ergebnis weist darauf hin, dass es verschiedene Routen der Verarbeitung gibt, die in Konflikt miteinander stehen. So steht auch die besondere Wirkmächtigkeit der Zeitzeug:innenberichte in Konflikt mit den Zielen des Geschichtsunterrichts, mit Quellen kritisch und reflektiert umzugehen. Für die geschichtsdidaktische Forschung zeigt die Dissertation folglich, dass das Lernen mit Zeitzeug:innenberichten differenziert betrachtet werden sollte und ein Blick auf die Lernprozesse der Lernenden (über das Lernmaterial und die Wirkungen hinaus) neue Erkenntnisse bietet.

Die Dissertationsstudien liefern zudem Einblicke in die Rolle von Emotionen beim historischen Lernen. Die geschichtsdidaktische Forschung beschäftigt sich seit Jahrzehnten mit der Frage, inwiefern Emotionen den Zugang zur Geschichte verschließen oder eröffnen können (Brauer & Lücke, 2013). Die Rolle von Emotionen im historischen Denkprozess ist mittlerweile anerkannt, allerdings sind die genauen Mechanismen noch nicht in Gänze verstanden und empirisch geprüft (Brauer & Lücke, 2013). Die Dissertation zeigt eindeutig, dass die Emotionen der Lernenden in Form der emotionalen Involviertheit mit Zeitzeug:innenberichten von besonderer Bedeutung sind. Denn die affektiven Verarbeitungsprozesse bilden (mit kognitiven Prozessen, aber nicht mit meta-kognitiven Prozessen) den Kern des distanzlos-involvierten Verarbeitungsmodus und evozieren das „Aura-Erlebnis“ mit Zeitzeug:innenberichten (Studie 1, Studie 2). Zudem zeigte sich, dass die Authentizität der Zeitzeug:innenberichte in Studie 3 insbesondere in Zusammenhang mit den affektiven Verarbeitungsprozessen steht und die distanzlos-involvierte Verarbeitung in negativem Zusammenhang zur distanziert-analytischen, also der reflektierten Verarbeitung der Zeitzeug:innenberichte. Betrachtet man also die Berichte ohne instruktionale Einbettung – wie dies in der vorliegenden Dissertation mit Ausnahme von Studie 4 der Fall war – lösten die Berichte einen distanzlos-involvierten Verarbeitungsmodus aus, der den distanziert-analytischen Verarbeitungsmodus erschwerte. (Studie 2 und Studie 3). Beide Modi, der distanziert-analytische sowie der gegenwartsbezogene Verarbeitungsmodus, wären mit Blick auf das Ziel des Geschichtsunterrichts erstrebenswert.



In Anlehnung an die vom Angebots-Nutzungs-Modell postulierte Bedeutung der „Wahrnehmung und Interpretation“ des Unterrichtsangebots für die Lernaktivität nimmt auch die vorliegende Dissertation ein zentrales Merkmal von Zeitzeug:innenberichten in den Blick: Die wahrgenommene Authentizität der Berichte. Bereits Angvik und von Borries (1997) sowie Rosenzweig und Thelen (1998) konnten in ihren Umfragen zeigen, dass Zeitzeug:innenberichte ein besonderes Vertrauen genießen – sogar ein größeres Vertrauen als Historiker:innen oder Lehrkräfte. Ballis und Schwendemann (2021) schlussfolgerten aus ihrer Studie, dass das Vertrauen den Berichten gegenüber das Fundament der Begegnungen zwischen Lernenden und Zeitzeug:innenberichten ist. Studie 2 und Studie 3 nahmen dieses Vertrauen über die wahrgenommene Authentizität differenziert in den Blick, um die zugrunde liegenden Lernprozesse besser zu verstehen. Den Studien lag in Anlehnung an das Angebots-Nutzungs-Modell die Annahme zugrunde, dass die wahrgenommene Authentizität des Lernmaterials den Effekt des Lernmaterials auf die Verarbeitungsprozesse mediiert. Studie 2 und 3 differenzierten dementsprechend das Konstrukt in verschiedene Facetten. Es fanden sich Hinweise, dass insbesondere die Facetten der Glaubwürdigkeit und der Lebendigkeit von Geschichte eine Rolle für das „Aura-Erlebnis“ spielen: Lernende attestierten den Zeitzeug:innenberichten eine besondere Glaubwürdigkeit und nahmen sie als außergewöhnliche Quellen wahr, um die Geschichte verständlich und lebendig zu vermitteln. Diese Wahrnehmungen standen in positivem Zusammenhang zum „Aura-Erlebnis“. Für die Annahme der vermittelten Effekte der Berichte auf die Verarbeitung, fanden die Studien konsistente Effekte für die Lebendigkeit von Geschichte, allerdings nicht für die Glaubwürdigkeit der Berichte. Diese Ergebnisse erfordern folglich weitere Studien zur Bedeutung der wahrgenommenen Authentizität als Mediator des Effekts von Lernmaterial auf Lernaktivität. Studie 3 nahm zudem die Authentizität auf einer weiteren Ebene in den Blick: die Authentizität der videografierten Zeitzeug:innenberichte in Form von unterschiedlichen Einführungen identischer Inhalte – als originaler Bericht, als nachgespielter Bericht und als rekonstruierter Bericht. Die Authentizität der Videoclips hatte einen Effekt auf die affektiven, aber nicht auf die kognitiven Verarbeitungsprozesse. Dieses Ergebnis ist insofern interessant, als lediglich das Framing der Videoclips variiert wurde und dies dennoch eine Bedeutung für die affektiven Verarbeitungsprozesse hatte. Die differenzierte Betrachtung der wahrgenommenen Authentizität zeigt, dass in Einklang mit bisherigen empirischen Studien (z. B. Angvik & von Borries, 1997; Ballis & Schwendemann, 2021; Rosenzweig & Thelen, 1998) dem Vertrauen, die Geschichte lebendig und verständlich zu machen, besondere Bedeutung zukommt. Für die geschichtsdidaktische Forschung zeigen sich also differenzierte Einblicke in die Konzeptualisierung und Erfassung der Authentizität als besondere Eigenschaft des

Lernarrangements. Damit bestätigen die Ergebnisse der Dissertation die Bedeutsamkeit des Vertrauens der Lernenden in die Berichte, Vergangenes verständlich zu machen.

Die Dissertation leistet allerdings nicht nur auf inhaltlicher Ebene einen Beitrag zu einem differenzierten Verständnis des Lernens mit Zeitzeug:innenberichten. Sie zeigt zudem durch die vielfältigen methodischen Zugänge neue empirische Wege für die Geschichtsdidaktik auf. Standardisierte Messinstrumente sind in geschichtsdidaktischer Forschung – die häufig auch qualitativ arbeitet – bisher nicht die Regel. Studie 1 bildete mit dem konzeptionellen Modell die Grundlage für die Entwicklung eines standardisierten Fragebogens. Denn ein theoretisches Verständnis des Forschungsgegenstands ist Voraussetzung für eine Überführung in ein standardisiertes Messinstrument. Darauf aufbauend pilotierte und validierte Studie 2 den Fragebogen, der aus Items bestehender Messinstrumente sowie aus selbst entwickelten und mit Expert:innen diskutierten Items bestand. Der validierte Fragebogen wurde schließlich genutzt, um in der experimentellen Online-Studie (Studie 3) Anwendung zu finden. Der entwickelte Fragebogen kann folglich in weiteren Studien eingesetzt werden. Studie 4 zeigt darüber hinaus die Möglichkeit von Sekundärdatenanalysen für die geschichtsdidaktische Forschung auf.

Zusammengefasst eröffnen die vier Dissertationsstudien ein tiefgreifendes Verständnis über das Lernen mit Zeitzeug:innenberichten, das über bisherige theoretische wie empirische Untersuchungen hinausgeht. Sie nehmen in Anlehnung an das Angebots-Nutzungs-Modell für unterrichtliche Wirkungsweisen verschiedene Einflussfaktoren im Rahmen der Lernprozesse in den Blick und tun dies mit vielfältigen methodischen Zugängen. Dabei zeigen sie auf, dass (a) die Lernaktivität in Form der situationalen Verarbeitung und des Involvements, (b) die affektiven Verarbeitungsprozesse der Lernenden und (c) die Wahrnehmung der Berichte als authentisch wichtige Faktoren beim Lernen mit Zeitzeug:innenberichten sind. Darüber hinaus bieten die Dissertationsstudien konzeptuelle Überlegungen zur besonderen Begegnung mit Zeitzeug:innenberichten (Studie 1 und Studie 2) und der wahrgenommenen Authentizität (Studie 2 und Studie 3) sowie entsprechende Messinstrumente an. Die Studien leisten allerdings nicht nur einen Beitrag zur Geschichtsdidaktik. Vielmehr führt die intensive Beschäftigung mit dem spezifischen Lernarrangement im Geschichtsunterricht auch zu Erkenntnissen für die Empirische Bildungsforschung, insbesondere für die allgemeine Unterrichtsforschung.

### **7.1.2 Lehr-Lernprozesse im (Fach-)Unterricht: Beitrag zur Empirischen Bildungs- und Unterrichtsforschung**

Das Lernarrangement mit Zeitzeug:innenberichten ist ein sehr spezifisches Lernarrangement im Geschichtsunterricht. Dennoch können die vorliegenden Studien wertvolle Erkenntnisse für die Empirische Bildungsforschung, im Speziellen für die pädagogisch-psychologische Unterrichtsforschung liefern. Seidel (2014) formulierte in ihrem Artikel zur Integration von Prozess- und Produkt-Paradigma in Angebots-Nutzungs-Modellen den Bedarf an theoretischen wie empirischen Ausschärfungen einzelner Aspekte des Angebots-Nutzungs-Modells durch interdisziplinäre Zusammenarbeit von Fachdidaktiken und generischer Unterrichtsforschung. Die vorliegenden Dissertationsstudien können genau zu dieser Ausdifferenzierung einen wesentlichen Beitrag leisten. Im Folgenden wird dieser Beitrag anhand von vier Punkten differenziert aufgeführt: (a) eine Ausdifferenzierung der Lernvoraussetzungen, (b) eine Ausdifferenzierung der Wahrnehmung und Interpretation, (c) eine Ausdifferenzierung der Lernaktivität und schließlich (d) eine Ausdifferenzierung der Lehr-Lernprozesse im Rahmen der Unterrichtsqualität.

Dass die Lernvoraussetzungen der Schüler:innen eine bedeutende Rolle in den Lernprozessen spielen, wurde mittlerweile vielfach empirisch bestätigt (z. B. OECD, 2005; Renkl, 2015; siehe Kapitel 1.3.1). Dabei untersuchten Studien als Einflussfaktoren häufig kognitive Lernvoraussetzungen wie das Vorwissen (z. B. Fyfe & Rittle-Johnson, 2016; Simonsmeier et al., 2022) oder motivational-affektive Voraussetzungen wie beispielsweise Motivation und Interesse (z. B. Hulleman et al., 2016; Schiefele et al., 1992). Mit zunehmender Heterogenität der Schüler:innenschaft wurde in den vergangenen Jahren die Forderung nach adaptivem Unterricht laut – also einem Unterricht, der die Voraussetzungen aller Lernenden berücksichtigt. Hinter dieser Forderung stehen die Annahme sowie Erkenntnisse darüber, dass das Unterrichtsangebot und die Lerngelegenheiten für Schüler:innen unterschiedlicher Voraussetzungen unterschiedlich wirken (z. B. Corno, 2008; Lehmann et al., 2016; Snow, 1991; Tetzlaff, Schmiedek & Brod, 2020). Die Studien, die im Rahmen der Aptitude-Treatment-Interaction-Forschung (Kapitel 1.3.1) generiert wurden, untersuchten bisher primär kognitive Lernvoraussetzungen und ihre Interaktion mit den Lehrmethoden auf die Wirkungen des Unterrichts – also weniger affektiv-motivationale Merkmale, obwohl diese in den theoretischen Ausführungen durchaus berücksichtigt wurden (Snow, 1992; Snow, Kanfer & Cudeck, 1989). Studie 4 dieser Dissertation berücksichtigte motivational-affektive Lernvoraussetzungen in Form des Wohlbefindens in sozialen Situationen und konnte Hinweise darauf finden, dass Zeitzeug:innenberichte in den Formaten Live-Begegnung vs. videografierte/verschriftlichte Berichte für Lernende mit

unterschiedlichem Wohlbefinden in sozialen Situationen unterschiedlich wirkten. Konkret: Lernende, die sich in sozialen Situationen eher unwohl fühlten, erzielten mit den video-grafierten und verschriftlichten Berichten höhere Ergebnisse in den Kompetenztests historischen Denkens. Lernende, die sich in sozialen Situationen wohlfühlten, erzielten in allen Bedingungen gleich gute Ergebnisse in den Kompetenztests. Die Ergebnisse der Studie können zum einen als Bestätigung der Ergebnisse im Rahmen der Forschung zu Aptitude-Treatment-Interaktionen gesehen werden. Insofern, als das die Lerngelegenheit insbesondere für Lernende mit schwächeren Lernvoraussetzungen eine Rolle spielt, für leistungsstärkere Lernende aber keinen Unterschied macht. Zum anderen gibt die Studie Hinweise darauf, dass die Heterogenität der Lernenden nicht nur in Bezug zu kognitiven Lernvoraussetzungen besteht, sondern auch in Bezug zu Merkmalen, die nicht direkt oder mittels Klassenarbeiten sichtbar sind. Die von Wylie und Chi (2014) entwickelte Klassifizierung von Lernaktivitäten legt nahe, dass insbesondere interaktive Tätigkeiten zu einem tiefgreifenden Verständnis des Lerngegenstandes führen (Kapitel 1.3.3). Folglich erfordert nicht nur der Unterricht mit Live-Zeitzeug:innen soziale Interaktion, sondern jede Unterrichtssituation, die von den Lernenden einen Umgang mit Mitschüler:innen oder der Lehrkraft verlangt. Deshalb erscheint es wichtig, auch in anderen Kontexten die Bedeutung des Wohlbefindens in sozialen Situationen zu untersuchen.

Die Dissertation differenziert das Verständnis der Wahrnehmung und Interpretation des Unterrichtsangebots aus: Das Angebots-Nutzungs-Modell nimmt an, dass das Unterrichtsangebot, vermittelt über die Wahrnehmung und Interpretation der Lernenden, auf die Lernaktivität wirkt. In den Studien der vorliegenden Arbeit wurde ein zentrales Merkmal des Lernens mit Zeitzeug:innenberichten im Rahmen der Wahrnehmung thematisiert: die wahrgenommene Authentizität. Diese ist allerdings nicht nur für das Lernen mit Zeitzeug:innenberichten bedeutsam, denn auch in anderen Fächern besteht der Anspruch, die Lernsituation so zu gestalten, dass sie Lernenden ein möglichst nahes Abbild von Situationen außerhalb der Schule gibt. Konkret heißt das, auch im naturwissenschaftlichen Unterricht werden durch Experimente und Besuche in Laboren authentische Lernarrangements geschaffen. Der bilinguale Unterricht etwa schafft durch die Verwendung der Sprache im Fachunterricht ein Lernarrangement, in dem die Sprache genutzt wird – und nicht nur gelernt. Im Ethikunterricht werden verschiedene Gotteshäuser besucht, um den Lernenden Einblicke in die verschiedenen Religionen und Kultstätten zu geben. Für das Lernen in authentischen Lernarrangements ist allerdings bisher nicht klar, inwiefern die Authentizität das Lernen tatsächlich fördert (z. B. Itzek-Greulich et al., 2015; Nachtigall & Rummel, 2021; Riegel & Kindermann, 2016). Zudem weisen Literatursynthesen auf unterschiedliche Facetten von

Authentizität hin (Fougt et al., 2019; Shaffer & Resnick, 1999), die eine differenzierte Betrachtung von authentischen Lernarrangements nahelegen. Die Ergebnisse der Dissertationsstudien zeigen auf, dass sich verschiedene Facetten von wahrgenommener Authentizität trennen lassen, und bestätigen damit die Ergebnisse von Hampp und Schwan (2014a, 2014b) und Hofer (2016). Die Ergebnisse zeigen aber auch, dass nicht alle Facetten gleichermaßen prädiktiv für die Verarbeitungsprozesse beim Lernen mit Zeitzeug:innenberichten sind. In Zusammenhang mit bisherigen Studien lässt sich daraus schließen, dass eine differenzierte Betrachtung von Authentizität im Hinblick auf das zu bewertende Subjekt oder Objekt (z. B. Zeitzeug:in, Videoclip, Lernmaterial) und das Lernergebnis einen wertvollen Erkenntnisgewinn liefert.

Die Dissertationsstudien können einen Beitrag zu einem differenzierten Verständnis der Lernaktivität leisten: Die vorliegenden Studien konzeptualisierten die Lernaktivität über mental ablaufende Verarbeitungsprozesse des Lernarrangements. Konkret stand dabei der distanzlos-involvierte Verarbeitungsmodus, ein Zusammenspiel aus geringer persönlicher Distanz zur Vergangenheit und niedrigem Abstraktionsniveau im Fokus der Studien 1 bis 3. In der Ausdifferenzierung der Verarbeitungsprozesse wird deutlich, dass hier sowohl affektive als auch kognitive Prozesse eine Rolle spielen. Kognitive Prozesse meinen hier allerdings nicht die kritische Auseinandersetzung mit der Thematik, sondern vielmehr ein kognitives Eintauchen in die Berichte der Zeitzeug:innen. Die Dissertationsstudien geben Hinweise darauf, dass sich die mental ablaufenden Verarbeitungsprozesse standardisiert erfassen lassen (Studie 2) und die distanzlos-involvierte Verarbeitung in negativem Zusammenhang mit der distanziert-analytischen Verarbeitung des Lernmaterials steht (Studie 2, Studie 3). Damit weisen die Studien auf die Bedeutung gegenstandsbezogener Emotionen im Lernkontext hin, die eine meta-kognitive Verarbeitung der Lerninhalte erschweren können. Das gilt nicht nur beim Lernen mit Zeitzeug:innenberichten. So ist beispielsweise denkbar, dass im Physikunterricht das Lernen über den Klimawandel mit Einzelbeispielen zu starken emotionalen Reaktionen der Lernenden führen, ein abstraktes Lernen über den Treibhauseffekt allerdings erschweren können. Mit Blick auf die persönlichen Handlungen der Lernenden stellt sich weiterführend die Frage, inwiefern die emotionalen Reaktionen auch fruchtbar sind. Ändere ich mein Mobilitäts- oder Essverhalten im Sinne des Klimas eher, wenn ich emotional betroffen bin oder den Klimawandel auf einer abstrakten Ebene verstehe, oder beides? Diese Fragen können möglicherweise über ein Verständnis der Beschaffenheit und Mechanismen des persönlich-motivierten Verarbeitungsmodus (Kapitel 1.3.3) beantwortet werden.

Schließlich werfen die Dissertationsstudien durch die Anwendung der Oberflächen- und Tiefenmerkmale auf ein konkretes Lernarrangement Fragen zu der Bedeutung dieser Merkmale

sowie zum Verständnis der Tiefenmerkmale auf. Das Modell der Basisdimensionen bildet die zentralen Lehr-Lernprozesse im Unterricht ab und verbindet im Angebots-Nutzungs-Modell das Unterrichtsangebot und die Lernaktivität (Kunter & Trautwein, 2013; Seidel, 2014). Trautwein et al. (2021) wiesen in ihrem Positionspapier auf den Gewinn einer Verknüpfung von Fachdidaktik Geschichte und allgemeiner Unterrichtsforschung hin und füllten die generischen Basisdimensionen mit fachdidaktischen Inhalten. Eine Verknüpfung der empirischen Studien zum Lernen mit Zeitzeug:innen mit den Oberflächen- und Tiefenmerkmalen konnte bereits einige Hinweise auf Gelingensfaktoren im Lernarrangement geben sowie die Notwendigkeit einer Ausdifferenzierung des Verständnisses der Basisdimensionen aufzeigen (Kapitel 1.2.4). Auch die vorliegenden Dissertationsstudien können einen Beitrag dazu leisten, die Oberflächen- und Tiefenmerkmale differenzierter zu verstehen.

Im Rahmen der Oberflächenmerkmale betrachteten die Dissertationsstudien sowohl videografierte Berichte (Studie 3, Studie 4) als auch Live-Begegnungen (Studie 2, Studie 4). Studie 2 und Studie 3 weisen darauf hin, dass die Verarbeitungsprozesse bei beiden Formaten mit dem standardisierten Fragebogen aus Studie 2 erfasst werden können. Das impliziert, dass das „Aura-Erlebnis“ sowohl von der Begegnung mit videografierten als auch mit Live-Zeitzeug:innenberichten ausgelöst werden kann. Allerdings gibt die Schwierigkeit der Erfassung der physischen Verarbeitungsprozesse in der Schüler:innenstichprobe (Studie 2) Hinweise darauf, dass das Format mit Blick auf diese Verarbeitungsprozesse doch eine zentrale Rolle spielen könnte. Auch Studie 4 zeichnet ein differenziertes Bild der Rolle des Formats auf das Lernen mit Zeitzeug:innenberichten und suggeriert eine Bedeutsamkeit der Oberflächenmerkmale.

Mit Blick auf die Tiefenmerkmale wird in den Studien (Studie 2, Studie 3) auf Ebene der Verarbeitungsprozesse zweierlei besonders deutlich: Die affektiven Verarbeitungsprozesse spielen eine zentrale Rolle und die kognitiven Verarbeitungsprozesse sind von meta-kognitiven Verarbeitungsprozessen, die sich eher durch Abstraktion und Transfer auszeichnen, abzugrenzen. Für die Dimension „Potenzial zur kognitiven Aktivierung“ stellt sich folglich die Frage, ob eine Form der affektiven Aktivierung nicht auch Berücksichtigung finden sollte. Zwar stehen die hier untersuchten affektiven Prozesse eher in negativem Zusammenhang zu einer analysierenden Verarbeitungsweise – die dem Ziel des Geschichtsunterrichts näher kommt. Allerdings erscheint es mit Blick auf das Erkennen der Relevanz des Geschichtsunterrichts notwendig, dass die Lernenden auch persönlich und emotional angesprochen werden (Van Straaten et al., 2016). Zudem wird durch die zwei betrachteten Verarbeitungsmodi deutlich (Kapitel

1.3.3, Studie 2, Studie 3), dass kognitive Aktivierung insbesondere eine Verarbeitungstiefe initiieren sollte, die einem hohen Abstraktionsniveau, also vielmehr einer meta-kognitiven Aktivierung, entspricht.

Zu dem Tiefenmerkmal der konstruktiven Unterstützung zeichnet der aktuelle Forschungsstand zum Lernen mit Zeitzeug:innenberichten ein konsistentes Bild: Die Berichte betreffen viele Schüler:innen emotional, insbesondere dann, wenn die Zeitzeug:innen Gewalt und Unrecht erfahren haben (Ballis & Schwendemann, 2021; Brüning, 2018; Obens & Geißler-Jagodzinski, 2010). Auch die vorliegenden Dissertationsstudien (Studie 2, Studie 3) bestätigen die emotionalen Anteile der Verarbeitung. Die emotionalen Verarbeitungsprozesse der Schüler:innen erfordern die sozio-emotionale Unterstützung durch die Lehrkraft, die Gefühle der Schüler:innen anerkennt und ihnen Raum gibt, die Gefühle zu thematisieren. Die konstruktive Unterstützung hat folglich einen stark sozio-emotionalen Anteil und bezieht sich nicht nur auf den Umgang der Lehrkraft mit den Lernenden. Zudem zeigt Studie 4 auf, dass Lernende in ihrer Unterschiedlichkeit auch mit Blick auf motivational-affektive Eigenschaften wahrgenommen werden sollten, um das Lernziel zu erreichen. In diesem Zusammenhang kann vermutet werden, dass die Begleitung durch die Lehrkraft eine zentrale Rolle dabei spielt, dass die affektiven Verarbeitungsprozesse die kritische Auseinandersetzung nicht behindern, sondern diese vielmehr als Katalysatoren für das Lernen dienen. Diese Überlegungen bedürfen allerdings weiterer Ausschärfungen sowie empirischer Prüfung.

Zusammengefasst zeigen die Dissertationsstudien den Erkenntnisgewinn für die Empirische Bildungsforschung anhand einer Ausdifferenzierung (a) der Lernvoraussetzungen, (b) der Wahrnehmung und Interpretation, (c) der Lernaktivität sowie (d) der Lehr-Lernprozesse im Rahmen der Oberflächen- und Tiefenmerkmale des Lernarrangements auf. Die Dissertation leistet einen wertvollen Beitrag zur Forschungslandschaft der Empirischen Bildungsforschung, die sich bisher wenig mit dem Fach Geschichte beschäftigte. Denn die systematische Untersuchung von Unterrichtsfächern wie Geschichte, aber auch angrenzenden Fächern wie Politik oder Sozialwissenschaften hat insbesondere mit Blick auf aktuelle gesellschaftliche Entwicklungen (siehe Eingangsbeispiele; Kapitel 1) eine bedeutende Relevanz: Für das friedliche Zusammenleben in einer Demokratie ist es wichtig, dass die Menschen in der Lage sind, sich begründete und kritische Urteile zu bilden und aus vergangenen Ereignissen für die Gegenwart und Zukunft zu lernen.

## 7.2 Stärken und Grenzen der Studien

Die Studien dieser Dissertation weisen Stärken und Limitationen auf, die es bei der Interpretation der Ergebnisse und der Einordnung der Befunde in einen größeren Forschungskontext zu berücksichtigen gilt.

Die Dissertation knüpft in ihrem Kern an ein zentrales Forschungsdesiderat aus der Geschichtsdidaktik an: Worin besteht die spezifische „Aura der Authentizität“ von Zeitzeug:innen und wie wirkt sich diese auf das Lernen aus? (Bertram, 2017; Brüning, 2018; Plato, 2001, Sabrow 2012). Sie folgte in diesem Zusammenhang konkret der Frage, welche Lernprozesse beim Lernen mit Zeitzeug:innenberichten bedeutend sind. Zur Beantwortung dieser Frage wurden vier Studien durchgeführt, die sowohl inhaltlich als auch methodisch vielfältige Zugänge bieten und somit zu einem systematischen und umfangreichen Verständnis des Lernens mit Zeitzeug:innenberichten beitragen. Insbesondere die unterschiedlichen empirischen Forschungsdesigns stellen eine Besonderheit für die geschichtsdidaktische Forschung dar (z. B. standardisiertes Messinstrument, Experiment). Zudem decken die Studien unterschiedliche Formate von Zeitzeug:innenberichten (Videoclips: Studie 3, Studie 4; Live: Studie 2, Studie 4), verschiedene Themen und historische Bezüge ab (Zweiter Weltkrieg: Studie 2 [Pilot], Studie 3; Alltagsgeschichten aus der ehemaligen DDR: Studie 3; Widerstand in der ehemaligen DDR: Studie 4; Haftenerfahrung in der ehemaligen DDR: Studie 2). Auch diese Vielfältigkeit trägt zu dem umfassenden Verständnis der Lernprozesse bei, welches durch die vorliegende Dissertation entwickelt wurde. Darüber hinaus schafft die Dissertation durch die fundierte Entwicklung eines standardisierten Messinstruments eine Grundlage zur weiteren Erforschung der Verarbeitungsprozesse im Rahmen des Lernens mit Zeitzeug:innenberichten und vergleichbaren Lernarrangements. Schließlich stand in der vorliegenden Dissertation das Lernen in einem Unterrichtsfach im Fokus, das bisher nur selten Forschungsgegenstand in der Empirischen Bildungsforschung war, obwohl die hier zu vermittelnden Kompetenzen für die schulischen Bildungsziele besonders zentral sind. Das generische Angebots-Nutzungs-Modell diente in der vorliegenden Dissertation als Rahmenmodell und ermöglichte so einen systematischen Zugang zur Erforschung des Lernens mit Zeitzeug:innenberichten. Als Beitrag zur Empirischen Bildungsforschung konnten in diesem Zusammenhang einzelne Bereiche des Angebots-Nutzungs-Modells wie beispielsweise die Lernaktivität ausdifferenziert werden (Kapitel 7.1.2). Den Bedarf an dieser Form der Ausdifferenzierung formulierte Seidel (2014) als Forschungsdesiderat interdisziplinärer Forschungsprojekte von Fachdidaktiken und Empirischer Bildungsforschung. Folglich zeigt die Dissertation in ihrer Gesamtschau den Erkenntnisgewinn interdisziplinärer Forschung zwischen Geschichtsdidaktik und Empirischer Bildungsforschung auf.



Die Studien der Dissertation weisen allerdings auch Limitationen auf, von denen ich drei im Folgenden diskutiere. Die erste Limitation umfasst die Stichprobe in den Studien, insbesondere den Schüler:innen- oder Studierenden-Status und das damit einhergehende Alter der Teilnehmenden. Ausgangspunkt der Dissertation waren die Ergebnisse der Studie von Bertram et al. (2017), die Schüler:innen am Ende der Sekundarstufe I untersuchte. Da Studie 4 mit dem gleichen Datensatz arbeitete, war die Stichprobe hier identisch: Schüler:innen der 9. Klasse im Geschichtsunterricht. Die anderen Studien wichen davon allerdings ab. Schon das konzeptuelle Modell in Studie 1 bezog sich nicht nur auf Lernende in schulischen Bildungskontexten, sondern umfasste auch Menschen, die historische Orte eigeninitiativ besuchen und auch Zeitzeug:innenberichten im Rahmen der Erinnerungskultur begegnen. Der standardisierte Fragebogen in Studie 2 wurde in Studierendenstichproben pilotiert und mit einer Stichprobe von Schüler:innen der Klassenstufen 8 bis 13 validiert. Die Stichprobe von Studie 3 umfasste Erwachsene, die mittels der Online-Plattform *prolific.org* akquiriert wurden und etwa zur Hälfte an Hochschulen eingeschrieben waren. Zwar fanden Rosenzweig und Thelen (1998) in ihrer Umfrage, dass das Vertrauen gegenüber Zeitzeug:innenberichten sich über weite Altersgruppen erstreckt. Allerdings ist sowohl den Zielen des Geschichtsunterrichts (z. B. Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016; Kapitel 1.1.2) als auch dem FUER-Kompetenz-Strukturmodell (Schreiber et al., 2006) zu entnehmen, dass historische Kompetenzen in verschiedene Kompetenzniveaus differenziert werden können: Am Ende der Schullaufbahn soll ein (selbst-)reflexives Geschichtsbewusstsein ausgebildet sein, mit dem man den Prozess historischen Denkens durchlaufen kann. Es stellt sich die Frage, in welchem Zusammenhang die Entwicklung historischer Kompetenzen mit dem Alter steht, und damit einhergehend, ob auch die Verarbeitungsprozesse altersspezifisch sind. Möglicherweise fehlt es jüngeren Schüler:innen an kognitiven Fähigkeiten, um Zeitzeug:innenberichte ausreichend kritisch reflektieren zu können. Dies impliziert die Annahme, dass auch das Alter einen Einfluss darauf hat, wie stark sich Lernende kognitiv von den Berichten distanzieren können.

Eine zweite Limitation der Studien bezieht sich auf die verwendeten Messinstrumente. In Studie 2 wurde ein standardisierter Fragebogen zur Erfassung der distanzlos-involvierten Verarbeitungsprozesse validiert. Diese Verarbeitungsprozesse empirisch zu erfassen, stand im Fokus von Studie 2 und Studie 3. Weniger Beachtung fand allerdings die Erfassung der anderen Verarbeitungsmodi, die ich im theoretischen Rahmen hergeleitet habe (Kapitel 1.3.3). Für den kritisch-distanzierten Modus der Verarbeitung nutzte ich in Studie 2 und Studie 3 das Instrument von Bertram (2015) zur Einsicht in die Perspektivität von Zeitzeug:innenberichten mit dem Argument, dass hier die kritische Auseinandersetzung mit den Berichten deutlich wird.

Allerdings bilden die Items weniger den Lernprozess, sondern eher ein Ergebnis des Umgangs mit den Berichten ab. Sie unterscheiden sich also im sprachlichen Bezug deutlich von den Items des in Studie 2 entwickelten Fragebogens. Darüber hinaus wurde die hier im Fokus stehende Lernaktivität nicht mit Kompetenzen historischen Denkens in Verbindung gebracht, d. h., die Wirkung der Lernaktivität auf den Ertrag, also auf die Kompetenzen historischen Denkens, bleibt in dieser Arbeit nur spekulativer Natur. Zudem wurden in den vorliegenden Studien keine validierten Instrumente der allgemeinen Unterrichtsforschung eingesetzt, obwohl dies mit Blick auf den Erkenntnisgewinn der interdisziplinären Forschung eine Vergleichbarkeit zu anderen Fächern und authentischen Lernarrangements hätte schaffen können.

Schließlich sind die Zeitzeug:innenberichte in den Studien der Dissertation bis auf Studie 4 nicht in einen instruktionalen Kontext eingebettet, der die Begegnung rahmt und möglicherweise Anregungen für eine kritische Auseinandersetzung gibt. Deshalb könnten Kritiker:innen argumentieren, dass der Konflikt zwischen Überwältigung durch die persönlichen Berichte und Ziel des Geschichtsunterrichts vor allem dann zustande kommt, wenn die Begegnung mit Zeitzeug:innenberichten weder vor- noch nachbereitet wird. Allerdings stellten Bertram et al. (2017) die besondere Wirkmächtigkeit der Live-Zeitzeug:innenberichte auch trotz einer Nachbereitungsstunde fest, in der die Lernenden verschiedene Zeitzeug:innenberichte miteinander verglichen, was Einsichten in die Multiperspektivität von Geschichte fördern sollte. Hier ist weitere Forschung notwendig, die der Frage nachgeht, wie die Einbettung der Begegnung gestaltet sein muss, damit sowohl die emotionale Involviertheit aufgefangen als auch die meta-kognitive Aktivierung gefördert werden kann – bestenfalls mit dem Ergebnis, Verarbeitungsprozesse anzuregen, die eine persönliche Orientierung ermöglichen. Im folgenden Kapitel arbeite ich Anknüpfungspunkte für weitere Forschung sowie Implikationen für die Praxis heraus.

## 7.3 Implikationen für Forschung und Praxis

Die Studien der vorliegenden Dissertation bieten Anknüpfungspunkte für zukünftige Forschung sowie für die außerschulische und schulische Bildungspraxis, welche ich im folgenden Kapitel aufzeige.

### 7.3.1 Implikationen und Anknüpfungspunkte für zukünftige Forschung

Die Dissertation wirft konzeptuelle, empirische wie methodische Fragen auf, die als Anknüpfungspunkte für weitere Forschung dienen können. Die Lernaktivität mit Zeitzeug:innenberichten wird in dieser Dissertation als mentaler Verarbeitungsprozess entlang der Dimensionen *Persönliche Distanz zur Vergangenheit* und *Abstraktionsniveau* konzeptualisiert (Kapitel 1.3.3). Insbesondere die Verarbeitung des Lernmaterials im distanzlos-involvierten Modus, dem „Aura-Erlebnis“, stand in den Dissertationsstudien im Vordergrund (Studie 1, 2 und 3). Die weiteren Verarbeitungsmodi, wie beispielsweise der distanziert-analytische oder der gegenwartsbezogene Modus der Verarbeitung, fanden keine umfassende Berücksichtigung in den Studien. Dabei ist es für das Verständnis des Lernarrangements von Bedeutung, welche Prozesse den anderen Verarbeitungsmodi zugrunde liegen und wie diese erfasst werden können. Mit Blick auf das Ziel des Geschichtsunterrichts, Kompetenzen historischen Denkens und insbesondere die Fähigkeit zu fördern, durch die Beschäftigung mit der Geschichte Orientierung in Gegenwart und Zukunft zu finden, erscheint insbesondere der Modus der gegenwartsbezogenen Verarbeitung von Bedeutung (geringe persönliche Distanz zur Vergangenheit, hohes Abstraktionsniveau). In Zusammenhang mit bisherigen Forschungsarbeiten zum Umgang mit den Berichten (z. B. Brüning, 2018; Hogervorst, 2020) sollte zukünftige Forschung den Blick über das „Aura-Erlebnis“ hinaus erweitern und auch die anderen möglichen Verarbeitungsmodi thematisieren. Die medienpsychologische Forschung gibt in Anlehnung an Vorderer (1992) beispielsweise Anknüpfungspunkte für das Verständnis und die Erfassung des distanziert-analytischen Verarbeitungsmodus. Studien zur Einschätzung der Relevanz des Geschichtsunterrichts von Van Straaten et al. (2016, 2018) können Anknüpfungspunkte für die Konzeption des gegenwartsbezogenen Modus der Verarbeitung liefern.

Darüber hinaus impliziert die vorliegende Dissertation, dass das Lernarrangement mit Zeitzeug:innenberichten in Form des „Aura-Erlebnisses“ kognitive, physische und vor allem affektive Prozesse auslöst (Studie 2, Studie 3). Auch beim Lernpotenzial spielen nicht nur kognitive Lernvoraussetzungen eine Rolle, sondern auch affektive (wie das Wohlbefinden in sozialen Situationen, Studie 4). Dass Emotionen auf vielfältige Weise eine Rolle in den

Unterrichts- und Lehr-Lernprozessen spielen, ist mittlerweile *Common Sense* in der Geschichtsdidaktik (z. B. Brauer & Lücke, 2013, 2022; Rösen, 2013) und auch der pädagogisch-psychologischen Unterrichtsforschung (z. B. Hascher & Brandenberger, 2018; Pekrun & Linnenbrink-Garcia, 2014). Inwiefern insbesondere die Emotionen mit Blick auf den Lerngegenstand – gegenstandsbezogene Emotionen – den Lernerfolg fördern oder behindern (Brauer & Lücke, 2013), ist bisher noch nicht ausreichend geklärt. Zukünftige Forschung sollte die Verschränkung von Emotionen und Kognition konzeptuell stärker ausarbeiten sowie empirisch prüfen. Deutlich wird das beispielsweise an dem Verständnis von Unterrichtsqualität anhand des Modells der drei Basisdimensionen (Klieme, 2001). In Zusammenhang mit den konzeptuellen Überlegungen von Trautwein et al. (2021) zeigt die vorliegende Dissertation auf, dass auch die affektive Aktivierung eine Rolle in den Lehr-Lernprozessen spielt, die – ohne eine konstruktive Unterstützung durch die Lehrkraft – möglicherweise die meta-kognitiven Lernprozesse behindern kann (Trautwein et al., 2021; Zülzdorf-Kersting, 2020; siehe auch Kapitel 1.2.4, Kapitel 7.1.2). Zukünftige Forschung sollte folglich jene Unterrichtssituationen in den Blick nehmen, die gegenstandsbezogene Emotionen auslösen, und im Rahmen dieser Situationen die Lehr-Lernprozesse anhand der Basisdimensionen erforschen. So kann das Verständnis für die Rolle von Emotionen in den Lehr-Lernprozessen, insbesondere mit Blick auf ein Zusammenspiel von kognitiver Aktivierung und konstruktiver Unterstützung, erweitert werden.

Mit diesen konzeptuellen Fragen ergeben sich wiederum Fragestellungen methodischer und empirischer Art. Studie 2 entwickelte einen standardisierten Fragebogen, der über die Selbstauskunft der Lernenden die kognitiven, affektiven und physischen Verarbeitungsprozesse abbildet. Allerdings zeigte sich in Studie 2 auch, dass die Items die physischen Verarbeitungsprozesse in der Schüler:innenstichprobe nicht sauber erfassen konnten. Grundsätzlich basieren die hier durchgeführten empirischen Studien auf standardisierten Fragebögen oder Kompetenztests. Zur Erforschung physischer oder auch affektiver Prozesse gibt es weitere vielversprechende Möglichkeiten: Beispielsweise können physische und affektive Reaktionen über den Hautwiderstand (z. B. Christopoulos, Uy & Yap, 2019; Sakaki, Ueno, Ponzio, Harley & Mather, 2019) oder mittels neurowissenschaftlicher Methoden (z. B. Szucs & Goswami, 2007) erfasst werden. Eine Verwendung dieser Erfassungsmethoden könnte noch differenziertere Auskünfte über die Rolle von Emotion und Physis, aber auch über ihr Zusammenspiel mit der (Meta-)Kognition in den betrachteten Lernprozessen geben. Darüber hinaus bieten neue digitale Technologien die Möglichkeit, weitere Forschungsfragen im Kontext des Lernens mit Zeitzeug:innenberichten zu beantworten. Beispielsweise können –

ähnlich der 3D-Hologramme von Zeitzeug:innen (Ballis & Gloe, 2020) – digitale Zeitzeugnisse erschaffen werden, mit welchen Lernende über eine Spracherkennungssoftware in Interaktion treten. So könnte in einem experimentellen Setting beispielsweise die Wirkung der Interaktionsmöglichkeiten mit den Zeitzeug:innen auf die Lernprozesse und den Kompetenzzuwachs erforscht werden.

Alle Studien der vorliegenden Dissertation nahmen das Lernarrangement mit Zeitzeug:innenberichten in den Blick. Allerdings berücksichtigte lediglich Studie 4 eine Einbettung der Begegnung mit den Berichten in eine Instruktion – die anderen Studien fokussierten sich stärker auf das Material bzw. das Medium an sich (siehe auch Kapitel 7.2). Brüning (2018) forderte in ihrer Studie eine besondere Form der Quellenkritik, die dem Medium der Zeitzeug:innenberichte gerecht werde. Zusammengenommen mit dem Überwältigungscharakter der Begegnungen mit Zeitzeug:innenberichten stellt sich folglich die Frage, wie Lernende im Umgang mit der Begegnung angeleitet werden müssen, damit sie nicht im „Aura-Erlebnis“ verhaftet bleiben, sondern bestenfalls zu Orientierung gelangen. Die Studie von Rodenhäuser (2012) gab Hinweise darauf, dass durch eine Anleitung die kritische Auseinandersetzung mit den Berichten gelingen kann. Zukünftige empirische Forschung sollte deshalb die instruktionale Einbettung der Begegnungen in den Blick nehmen und untersuchen, welche Einbettung hinsichtlich des Lernergebnisses wirksam ist. Vielversprechende Anknüpfungspunkte hierfür bieten Studien zur Initiierung von *conceptual change* (z. B. Schneider & Hardy, 2013; Vosniadou, 1994; Vosniadou & Mason, 2012). Denn die bei Lernenden verbreitete Vorstellung, Zeitzeug:innen können *wahre* Einblicke in die Vergangenheit geben, erfordert ein Umdenken. Inwiefern die Instruktion in diesem Lernarrangement gestaltet sein muss, um im Sinne eines *conceptual change* wirksam zu sein, ist ein Forschungsdesiderat für zukünftige Studien.

Eine weitere Implikation für empirische Forschung ergibt sich aus den in Studie 2, 3 und 4 berechneten Mediationsmodellen. Die Studien zeigten, dass die wahrgenommene Authentizität den Effekt der Begegnung (Studie 2) bzw. des Lernmaterials (Studie 3) auf die Lernaktivität mediiert, so wie auch vom Angebots-Nutzungs-Modell angenommen. Auch Studie 4 untersuchte eine Mediationshypothese, fand für diese allerdings keine Bestätigung. Dennoch weisen insbesondere Studie 2 und Studie 3 auf die Bedeutung von Mediationsmodellen hin und implizieren für die weitere Forschung, dass diese im Kontext der Lernprozesse stärker untersucht werden sollten. Mit Blick auf das hier untersuchte Lernarrangement stellt sich insbesondere die Frage, ob die Lernaktivität in Form der Verarbeitung den Effekt der Begegnung mit Zeitzeug:innenberichten auf die Wirkung mediiert. Dieser Zusammenhang wird in der vorliegenden Dissertation angenommen, ist bisher allerdings nicht empirisch untersucht.

Hierzu sollte zukünftige Forschung die Lernaktivität – zum Beispiel durch den in Studie 2 entwickelten Fragebogen – als Mediator der Wirkung von Zeitzeug:innenberichten auf Kompetenzen historischen Denkens (gemessen beispielsweise durch standardisierte Kompetenztests zum historischen Denken, z. B. Trautwein et al., 2017) untersuchen.

### **7.3.2 Implikationen für die Praxis**

Ziel der vorliegenden Dissertation war es, das Lernarrangement mit Zeitzeug:innenberichten durch die Perspektiven von Geschichtsdidaktik und Empirischer Bildungsforschung systematisch und differenziert zu verstehen, um so Erklärungsansätze für den Konflikt zwischen Wirkungen des Lernarrangements und Zielen des Geschichtsunterrichts zu finden. Die Studien der Dissertation sind deshalb primär so konzipiert, dass sie einen Beitrag für die Forschung leisten. Es lassen sich auch Implikationen für die schulische wie außerschulische Bildungspraxis ableiten, die allerdings stets mit Blick auf die Studiendesigns und Stichproben betrachtet werden sollten.

Die Begegnung mit Zeitzeug:innenberichten löst kognitive, affektive und physische Verarbeitungsprozesse in den Lernenden aus. Diese Verarbeitungsprozesse lassen die Lernenden in die Vergangenheit eintauchen. Für diese Verarbeitung spielt auch die Wahrnehmung der Berichte als authentisch eine bedeutsame Rolle. Die Begegnung mit den Berichten ist folglich komplex. Lehrkräfte sollten sich bei der Gestaltung des Unterrichts mit Zeitzeug:innenberichten dieser Komplexität bewusst werden und den Lernenden Raum geben, ihre Emotionen und ihre individuellen Erfahrungen mit den Berichten zu thematisieren. Der Fragebogen aus Studie 2 könnte beispielsweise in reduzierter Form für Gesprächsanreize in einer Reflexionsstunde, die auch der Wahrnehmung der Lernenden gerecht wird, genutzt werden. Zudem zeigt Studie 4 (siehe auch Kapitel 1.3.1) die Bedeutung der Lernvoraussetzungen für das Lernen mit Zeitzeug:innenberichten auf. Lehrkräfte stehen hier der besonderen Herausforderung gegenüber, sensibel auf die Lernenden in der Gruppe zu reagieren. Sie sollten in der Lage sein – womöglich über Prozesse der konstruktiven Unterstützung (siehe Kapitel 1.2.4 und Kapitel 7.2) – jene Schüler:innen aufzufangen, die entweder zu schüchtern sind, um sich aktiv zu beteiligen (Studie 4), oder aufgrund ihrer Herkunft einen persönlichen Bezug zur Thematik haben (Kapitel 1.3.1). Die negativen Zusammenhänge zwischen distanzlos-involvierter und distanziert-analytischer Verarbeitung (Studie 2) können auch Lehrkräften Hinweise über die Risiken des Einsatzes der Zeitzeug:innenberichte geben und dafür sensibilisieren, die Zeitzeug:innenberichte mit klarem didaktischen Ziel einzusetzen und den Unterricht so zu gestalten, dass der Einsatz diesem Ziel gerecht wird.

Ein Verständnis über die Komplexität der Lernprozesse ist auch für Akteur:innen der außerschulischen Bildung bedeutsam. Denn insbesondere neuere Formate der Erinnerungskultur wie beispielsweise die Augmented-Reality-App des WDR oder das Instagram-Projekt des SWR (siehe Kapitel 1.2.2) bieten bisher kaum Reflexionsmöglichkeiten für die Nutzer:innen an. Dass dadurch möglicherweise Fehlvorstellungen von Geschichte gefördert werden, sollte den Produzent:innen dieser Formate bewusst sein. Bestenfalls entwickeln die Produzent:innen – wenn diese für sich einen Bildungsauftrag definieren – Möglichkeiten, auch die kritische Auseinandersetzung mit den Formaten und Perspektiven zu initiieren. Das erscheint mit Blick auf die Abhängigkeit von hohen Nutzer:innenzahlen schwierig, ist aber für ein Format mit Bildungsauftrag notwendig. Auch in „klassischen“ außerschulischen Bildungskontexten kann ein Verständnis über die Lernprozesse hilfreich sein. Studie 1 berücksichtigte explizit auch historische Orte in der Entwicklung des konzeptuellen Modells: Museen und Gedenkstätten arbeiten häufig mit emotionalen Zugängen, nutzen authentische Objekte oder Zeitzeug:innenberichte, um bestimmte Geschichten zu erzählen. Wie die authentischen Zugänge auf die Besucher:innen wirken, sollte von den Kurator:innen auch empirisch untersucht werden. So könnte im Hinblick auf die verschiedenen Verarbeitungsmodi eine Ausstellung gestaltet werden, die den emotionalen Zugang durch das Eintauchen in die Geschichten zwar nutzt, jedoch Lernende mittels Kontextualisierung dahingehend begleitet, dass sie nicht in den Geschichten verbleiben, sondern diese am Ende der Ausstellung reflektiert einordnen können. Schließlich zeigte Studie 3, die den Effekt unterschiedlicher Einführungen von Zeitzeug:innenberichten in TV-Dokumentationen untersuchte, dass die Einführung eine bedeutende Rolle für die emotionale Involviertheit hat. Für Ausstellungen mit videografierten Zeitzeugnissen oder auch für die Produktion von Filmmaterial sollte also auch durchdacht sein, wie die Berichte gerahmt sind, welche Aussage sie treffen, welche Informationen den Rezipient:innen vorab gegeben werden.

Zusammengefasst zeigt die Dissertation auch Akteur:innen aus der Bildungspraxis ein differenziertes Verständnis des Lernens mit Zeitzeug:innen auf. Die unterschiedlichen Verarbeitungsmodi geben Anknüpfungspunkte, die Begegnungen so zu gestalten, dass auch eine kritische Reflexion für die Lernenden möglich wird und bestenfalls eine persönliche Orientierung in Gegenwart und Zukunft stattfindet.

## 7.4 Fazit

Die vorliegende Dissertation erweitert das bisherige Verständnis über das Lernen mit Zeitzeug:innenberichten durch die systematische Untersuchung von Lernprozessen durch eine Verknüpfung von Geschichtsdidaktik und Empirischer Bildungsforschung. Die Dissertation konzeptualisiert die Verarbeitungsprozesse von Lernenden im Lernarrangement mit Zeitzeug:innenberichten und nimmt dabei insbesondere das „Aura-Erlebnis“ in den Blick. Dieses wird durch die vorliegende Dissertation in einem standardisierten Fragebogen messbar gemacht und in Zusammenhang mit der Authentizität des Lernmaterials sowie der Wahrnehmung der Lernenden gebracht. Zudem zeigt die Dissertation auf, dass eine Erweiterung des Verständnisses von Lernvoraussetzungen wertvolle Erkenntnisse zur differenziellen Wirksamkeit des Lernarrangements mit Zeitzeug:innenberichten liefert. In der Dissertation wird deutlich, dass eine empirisch forschende Geschichtsdidaktik, die sich an den Standards der Empirischen Bildungsforschung orientiert, und eine Empirische Bildungsforschung, die sich mit Lernarrangements in Fächern der historisch-politischen Bildung beschäftigt, wertvolle Erkenntnisgewinne erzielen können. Die interdisziplinäre Arbeit ist gerade deshalb von großer Relevanz, weil die Ziele des Geschichtsunterrichts auch das umfassen, was im Kern das Ziel moderner Schulbildung ist und so wichtig für das friedliche Zusammenleben in einer Demokratie: dass Schüler:innen lernen, eigenständig und reflektiert zu denken und zu urteilen.



---

# 8

## LITERATUR

---

## 8 Literatur

- Angvik, M. & von Borries, B. (1997). *Youth and history. A comparative European survey on historical consciousness and political attitudes among adolescents*. Hamburg: Körber-Stiftung.
- Ballis, A. & Gloe, M. (2019). *Holocaust Education Revisited*. Wiesbaden: Springer.
- Ballis, A. & Gloe, M. (2020). Interactive 3D testimonies of Holocaust survivors in German language. In A. Ballis & M. Gloe (Hrsg.), *Holocaust education revisited* (S. 343–368). Wiesbaden: Springer. [https://doi.org/10.1007/978-3-658-24207-7\\_21](https://doi.org/10.1007/978-3-658-24207-7_21)
- Ballis, A. & Schwendemann, L. (2021). 'In any case, you believe him one hundred percent, everything he says.' Trustworthiness in Holocaust survivor talks with high school students in Germany. *Holocaust Studies*, 28(2), 191–220. <https://doi.org/10.1080/17504902.2021.1915016>
- Baron, C. (2012). Understanding historical thinking at historic sites. *Journal of Educational Psychology*, 104(3), 833–847. <https://doi.org/10.1037/a0027476>
- Barricelli, M. & Sauer, M. (2006). Was ist guter Geschichtsunterricht? *Geschichte in Wissenschaft und Unterricht*, 4–26.
- Baumert, J. & Klieme, E. (2001). Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn - Die Herausforderung von TIMSS für die Weiterentwicklung des mathematischen und naturwissenschaftlichen Unterrichts. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS - Impulse für Schule und Unterricht* (S. 11–42). BMBF.
- Bernstein, J. (2021). *Antisemitismus an Schulen in Deutschland: Befunde - Analysen - Handlungsoptionen*. Bonn: Bundeszentrale für politische Bildung.
- Bertram, C. (2017). *Zeitzeugen im Geschichtsunterricht. Chance oder Risiko für historisches Lernen? Eine randomisierte Interventionsstudie*. Schwalbach/Ts.: Wochenschau.
- Bertram, C. (2020). Live oder medial vermittelt: Was lernen Schülerinnen und Schüler mit Zeitzeuginnen und Zeitzeugen? In A.-B. Rothstein & S. Pilzweger-Steiner (Hrsg.), *Entgrenzte Erinnerung* (S. 247–266). Oldenburg: De Gruyter.
- Bertram, C., Wagner, W. & Trautwein, U. (2014). Zeitzeugenbefragungen im Geschichtsunterricht: Entwicklung eines Kurzinstruments für die Wirksamkeitsmessung. In T. Arand & M. Seidenfuß (Hrsg.), *Neue Wege - neue Themen - neue Methoden? Ein Querschnitt aus der geschichtsdidaktischen Forschung des wissenschaftlichen Nachwuchses* (S. 191–208). Göttingen: V&R Academic.
- Bertram, C., Wagner, W. & Trautwein, U. (2017). Learning historical thinking with oral

- history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54(3), 444–484. <https://doi.org/10.3102/0002831217694833>
- Bilandzic, H. & Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, 36(1), 29–50. <https://doi.org/10.1515/COMM.2011.002>
- Bilandzic, H., Schnell, C., Sukalla, F., Hastall, M. R. & Busselle, R. W. (2019). The narrative engageability scale: A multidimensional trait measure for the propensity to become engaged in a story. *International Journal of Communication*, 13, 801–832.
- von Borries, B. (1988). *Geschichtslernen und Geschichtsbewußtsein. Empirische Erkundungen zu Erwerb und Gebrauch von Historie*. Stuttgart: Klett.
- von Borries, B. (1994). Von gesinnungsbildenden Erlebnissen zur Kultivierung der Affekte? Über die Ziele und Wirkungen von Geschichtslernen. *Studien zur internationalen Schulbuchforschung*, 76, 67–92.
- Bracke, S. & Flaving, C. (2018). Emotionen im Geschichtsunterricht. In S. Bracke, C. Flaving, J. Jansen, M. Köster, J. Lahmer-Gebauer, S. Lankes et al. (Hrsg.), *Theorie des Geschichtsunterrichts*. Schwalbach/Ts.: Wochenschau.
- Bracke, S., Flaving, C., Jansen, J., Köster, M., Lahmer-Gebauer, J., Lankes, S. et al. (2018). *Theorie des Geschichtsunterrichts*. Schwalbach/Ts.: Wochenschau.
- Brauer, J. (2016). „Heiße Geschichte“? Emotionen und historisches Lernen in Museen und Gedenkstätten. In S. Willner, G. Koch & S. Samida (Hrsg.), *Doing History - Performative Praktiken in der Geschichtskultur* (S. 29–44). Münster: Waxmann.
- Brauer, J. (2017, Oktober). Zeitzeugenbegegnung als Erlebnis im Geschichtsunterricht: Boom der Zeitzeugenschaft. *Blog der Heidelberg School of Education*. Verfügbar unter: <https://hse.hypotheses.org/668>
- Brauer, J. (2019). Gefühlte Geschichte? Emotionen, Geschichte und historisches Lernen. *Vierteljahresschrift für wissenschaftliche Pädagogik*, 95, 272–283. <https://doi.org/10.30965/25890581-09501022>
- Brauer, J. (2020, September 5). Kritisches Nachdenken über Zeugen der Erinnerung. *Deutschlandfunk Kultur*. Verfügbar unter: <https://www.deutschlandfunkkultur.de/gedenkstaette-berlin-hohenschoenhausen-kritisches-100.html>
- Brauer, J. & Lücke, M. (Hrsg.) (2013). *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven*. Göttingen: V&R

- unipress.
- Brauer, J. & Lücke, M. (2022). Emotionen, Geschichte und historisches Lernen. In M. Gläser-Zikuda, F. Hofmann & V. Frederking (Hrsg.), *Emotionen im Unterricht. Psychologische, pädagogische und fachdidaktische Perspektiven* (S. 168–178). Stuttgart: Kohlhammer.
- Brooks, S. (2009). Historical empathy in the social studies classroom: A review of the literature. *The Journal of Social Studies Research*, 33(2), 71–91.
- Brühwiler, C. & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21, 95–108.
- Brüning, C. I. (2018). *Holocaust Education in der heterogenen Gesellschaft*. Frankfurt am Main: Wochenschau.
- Bundesstiftung zur Aufarbeitung der SED-Diktatur (2016). *Gelebte Geschichte. DDR-Zeitzeugen in Schulen. Ein Leitfaden für Lehrkräfte*. Berlin: Bundesstiftung zur Aufarbeitung der SED-Diktatur.
- Busselle, R. & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, 18(2), 255–280. <https://doi.org/10.1111/j.1468-2885.2008.00322.x>
- Chi, M. T. H. & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Christopoulos, G. I., Uy, M. A. & Yap, W. J. (2019). The body and the brain: Measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods*, 22(1), 394–420. <https://doi.org/10.1177/1094428116681073>
- Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, 43(3), 161–173. <https://doi.org/10.1080/00461520802178466>
- Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12(11), 671–684. <https://doi.org/10.1037/h0043943>
- Csikszentmihalyi, M. (1991). *Flow*. New York: Harper & Row.
- Decristan, J., Hess, M., Holzberger, D. & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale. *Zeitschrift für Pädagogik, Beiheft*, 66(1), 102–116.
- Deutsche Presse Agentur (dpa) (2020, November 24). Antisemitismusbeauftragter rügt „krude Verharmlosungen“ des Holocaust. *Der Tagesspiegel*.
- Dittmer, N. (2019). Gute App - aber bitte nicht überwältigen lassen. *Deutschlandfunk Kultur*.

- Verfügbar unter: <https://www.deutschlandfunkkultur.de/wdr-app-zur-ns-zeit-gute-app-aber-bitte-nicht-100.html>
- Van Drie, J. & Van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, 20(2), 87–110. <https://doi.org/10.1007/s10648-007-9056-1>
- Dunkin, M. & Biddle, B. (1974). *The study of teaching*. New York: Holt, Rhinehart & Winston.
- Dutt-Doner, K. M., Allen, S. & Campanaro, K. (2016). Understanding the impact of using oral histories in the classroom. *The Social Studies*, 107(6), 257–265. <https://doi.org/10.1080/00377996.2016.1221792>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040–1048. <https://doi.org/10.1037/0003-066X.41.10.1040>
- Endacott, J. (2010). Reconsidering affective engagement in historical empathy. *Theory and Research in Social Education*, 38(1), 6–47. <https://doi.org/10.1080/00933104.2010.10473415>
- Endacott, J. & Brooks, S. (2013). An updated theoretical and practical model for promoting historical empathy. *Social Studies Research and Practice*, 8(1), 41–58.
- Endacott, J. & Brooks, S. (2018). Historical empathy: Perspectives and Responding to the past. In S.A. Metzger & L. McArthur Harris (Hrsg.), *The Wiley International Handbook of History Teaching and Learning* (S. 203–225). New York: Wiley-Blackwell.
- Erbar, R. (2012). Zeugen der Zeit? Zeitzeugengespräche in Wissenschaft und Unterricht. *Geschichte für heute*, 3, 5–21.
- Erl, A. (2011). *Kollektives Gedächtnis und Erinnerungskulturen*. Wiesbaden: Springer.
- Ernst, C. (2012). Zeitzeugenarbeit zur DDR-Geschichte. Historische Entwicklungslinien - Konzepte - Bildungspraxis. *Werkhefte für politische Bildung*.
- Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*. Wiesbaden: Springer.
- Flunger, B., Mayer, A. & Umbach, N. (2019). Beneficial for some or for everyone? Exploring the effects of an autonomy-supportive intervention in the real-life classroom. *Journal of Educational Psychology*, 111(2), 210–234. <https://doi.org/http://dx.doi.org/10.1037/edu0000284>
- Fougt, S. S., Misfeldt, M. & Shaffer, D. W. (2019). Realistic authenticity. *Journal of Interactive Learning Research*, 30(4), 477–504.

- Frazier, B. N., Gelman, S. A., Wilson, A. & Hood, B. M. (2009). Picasso paintings, moon rocks, and hand-written beatles lyrics: Adults' evaluations of authentic objects. *Journal of Cognition and Culture*, 9, 1–14.
- Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109.  
<https://doi.org/10.3102/00346543074001059>
- Frevert, U. (2011). *Emotions in history: lost and found*. Budapest: Central European University Press.
- Frevert, U. (2018). Politische Bildung - mit Gefühl? *Politik und Zeitgeschichte*, 68(13–14), 18–24.
- Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A. et al. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, 106(2), 499–514.  
<https://doi.org/10.1037/a0034341>
- Fulgini, A. J. (1997). The academic achievement of adolescents from immigrant families: The roles of family background, attitudes, and behavior. *Child Development*, 68(2), 351–363.
- Fyfe, E. R. & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, 108(1), 82–97.  
<https://doi.org/10.1037/edu0000053>
- Galda, M. (2013). *Geschichtsbewusstsein, historisches Wissen und Interesse*. Dissertation. Johann-Wolfgang-Goethe-Universität zu Frankfurt am Main.
- Gaspard, H., Dicke, A., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B. et al. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental psychology*, 51(9), 1226–1240.  
<https://doi.org/10.1037/dev0000028>
- Gautschi, P. (2009). *Guter Geschichtsunterricht*. Schwalbach/Ts.: Wochenschau.
- Gautschi, P., Hodel, J., Utz, H. & August, F. (2009). *Kompetenzmodell für „Historisches Lernen“ – eine Orientierungshilfe für Lehrerinnen und Lehrer*. Pädagogische Hochschule der Fachhochschule Nordwestschweiz. Verfügbar unter: [https://ernst-goebel.hoehchst.schule.hessen.de/fach/geschichte/material\\_geschichte/allpaed\\_geschichte/kompetenzorientierunggu/litkompetenzorientierunggu/Gautschi-Kompetenzmodell\\_fuer\\_historisches\\_LernenAug09.pdf](https://ernst-goebel.hoehchst.schule.hessen.de/fach/geschichte/material_geschichte/allpaed_geschichte/kompetenzorientierunggu/litkompetenzorientierunggu/Gautschi-Kompetenzmodell_fuer_historisches_LernenAug09.pdf)
- Goetz, T., Zirngibl, A., Pekrun, R. & Hall, N. (2003). Emotions, learning and achievement

- from an educational-psychological perspective. In P. Mayring & C. von Rhoeneck (Hrsg.), *Learning emotions: The influence of affective factors on classroom learning* (S. 9–28). Frankfurt am Main: Peter Lang.
- Goldberg, T. (2014). Looking at their side of the conflict? Effects of single versus multiple perspective history teaching on Jewish and Arab adolescents' attitude to out-group narratives and in-group responsibility. *Intercultural Education*, 25(6), 453–467.  
<https://doi.org/10.1080/14675986.2014.990230>
- Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, 38(2), 247–266.  
<https://doi.org/10.1207/s15326950dp3802>
- Green, M. C. & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721.  
<https://doi.org/10.1037/0022-3514.79.5.701>
- Gudjons, H., Wagener-Gudjons, B. & Pieper, M. (2008). *Auf meinen Spuren: Übungen zur Biografiearbeit*. Leipzig: Klinkhardt.
- Guignon, C. (2008). Authenticity. *Philosophy Compass*, 2, 277–290.
- Häfner, I., Flunger, B., Dicke, A. L., Gaspard, H., Brisson, B. M., Nagengast, B. et al. (2018). The role of family characteristics for students' academic outcomes: A person-centered approach. *Child Development*, 89(4), 1405–1422. <https://doi.org/10.1111/cdev.12809>
- Hampp, C. & Schwan, S. (2014a). Perception and evaluation of authentic objects: Findings from a visitor study. *Museum Management and Curatorship*, 29(4), 349–367.  
<https://doi.org/10.1080/09647775.2014.938416>
- Hampp, C. & Schwan, S. (2014b). The role of authentic objects in museums of the history of science and technology: Findings from a visitor study. *International Journal of Science Education, Part B*, 5(2), 161–181.  
<https://doi.org/10.1080/21548455.2013.875238>
- Hampp, C. & Schwan, S. (2017). Authentizität in der Wahrnehmung und Bewertung von Museumsobjekten. In T. Eser, M. Farrenkopf, M. Kimmel, A. Saupe & U. Warnke (Hrsg.), *Authentisierung im Museum: Ein Werkstatt-Bericht (RGZM – Tagungen, Band 32)* (S. 89–100). Mainz: Verlag des Römisch-Germanischen Zentralmuseums.
- Handler, R. (1986). Authenticity. *Anthropology Today*, 2(1), 2–4.
- Handler, R. & Saxton, W. (1988). Dyssimulation: Reflexivity, narrative and the quest for authenticity in „living history“. *Cultural Anthropology*, 3(3), 242–260.  
<https://doi.org/10.1525/can.1988.3.3.02a00020>

- Harackiewicz, J. M., Smith, J. L. & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences*, 3(2), 220–227. <https://doi.org/10.1177/2372732216655542>
- Hasberg, W. (2005). Von PISA nach Berlin. *Geschichte in Wissenschaft und Unterricht*, 56, 684–702.
- Hasberg, W. (2013). Emotionalität historischen Lernens. Einblicke in und Ausblicke auf empirische Forschung. In J. Brauer & M. Lücke (Hrsg.), *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven* (S. 47–74). Göttingen: V&R unipress.
- Hasberg, W. & Körber, A. (2003). Geschichtsbewusstsein dynamisch. In A. Körber (Hrsg.), *Geschichte - Leben - Lernen. Bodo von Borries zum 60. Geburtstag* (S. 177–200). Schwalbach/Ts.: Wochenschau.
- Hascher, T. & Brandenberger, C. (2018). Emotionen und Lernen im Unterricht. In M. Huber & S. Krause (Hrsg.), *Bildung und Emotion* (S. 289–313). Wiesbaden: Springer.
- Hattie, J. (2009). *Visible Learning—A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Haydn, T. & Harris, R. (2010). Pupil perspectives on the purposes and benefits of studying history in high school: A view from the UK. *Journal of Curriculum Studies*, 42(2), 241–261. <https://doi.org/10.1080/00220270903403189>
- Heafner, T. L. & Fitchett, P. G. (2015). An opportunity to learn US history: What NAEP data suggest regarding the opportunity gap. *The High School Journal*, 98(3), 226–249.
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Stuttgart: Klett-Kallmeyer.
- Helmke, A. & Weinert, F. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Enzyklopädie der Psychologie* (Bd. 3., S. 71–176). Göttingen: Hogrefe.
- Helmsing, M. (2014). Virtuous subjects: A critical analysis of the affective substance of social studies education. *Theory and Research in Social Education*, 42(1), 127–140. <https://doi.org/10.1080/00933104.2013.842530>
- Henig, L. & Ebbrecht-Hartmann, T. (2022). Witnessing Eva Stories: Media witnessing and self-inscription in social media memory. *New Media & Society*, 24(1), 202–226. <https://doi.org/10.1177/1461444820963805>
- Henke-Bockschatz, G. (2007). Zeitzeugenbefragung. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 354–369).



- Schwalbach/Ts.: Wochenschau.
- Henke-Bockschatz, G. (2014). *Oral History im Geschichtsunterricht*. Schwalbach/Ts.: Wochenschau.
- Herzing, J. & Longerich, M. (2013, Januar 26). Wenn die Zeitzeugen gehen. *Deutschlandfunk.de*. Verfügbar unter: <https://www.deutschlandfunk.de/wenn-die-zeitzeugen-gehen-100.html>
- Heyl, M. (1995). Betroffenheit ist kein Lernziel. Ein schwieriges Thema: Der Holocaust im Unterricht. *Erziehung und Wissenschaft*, 46(4), 14–16.
- Higgins, S. & Katsipataki, M. (2015). Evidence from meta-analysis about parental involvement in education which supports their children's learning. *Journal of Children's Services*, 10(3), 280–290. <https://doi.org/10.1108/JCS-02-2015-0009>
- Hofer, M. (2016). Der Einfluss von Faktualität und Fiktionalität auf eudaimonisches Unterhaltungserleben – die mediierende Rolle wahrgenommener Realität. In M. Czichon, C. Wunsch & M. Dohle (Hrsg.), *Rezeption und Wirkung fiktionaler Medieninhalte* (S. 79–96). Baden-Baden: Nomos.
- Hogervorst, S. (2018). Distanced by the screen. Student history teachers and video archives of Second World War interviews in the Netherlands. In W. Dreier, A. Laumer & M. Wein (Hrsg.), *Explorations of good practice in educational work with video testimonies of victims of national socialism* (S. 145–152). Berlin: Stiftung „Erinnerung, Verantwortung und Zukunft“.
- Hogervorst, S. (2020). The era of the user. Testimonies in the digital age. *Rethinking History*, 24(2), 169–183. Routledge. <https://doi.org/10.1080/13642529.2020.1757333>
- Hulleman, C. S., Barron, K. E., Kosovich, J. J. & Lazowski, R. A. (2016). Student motivation: Current theories, constructs, and interventions within an expectancy-value framework. In A. A. Lipnevich, F. Preckel & R. D. Roberts (Hrsg.), *Psychosocial skills and school systems in the 21st century* (S. 241–278). Cham: Springer. [https://doi.org/10.1007/978-3-319-28606-8\\_10](https://doi.org/10.1007/978-3-319-28606-8_10)
- Hulleman, C. S. & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326(December), 1410–1412.
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M. & Trautwein, U. (2015). Effects of a science center outreach lab on school students' achievement. *Learning and Instruction*, 38, 43–52. <https://doi.org/10.1016/j.learninstruc.2015.03.003>
- Jeismann, K.-E. (1977). Didaktik der Geschichte. In E. Kosthorst (Hrsg.),

- Geschichtswissenschaft - Didaktik - Forschung - Theorie* (S. 9–33). Göttingen: Vandenhoeck & Ruprecht.
- Jeismann, K.-E. (1994). Emotionen und historisches Lernen. Bemerkungen zur Tagung der Konferenz für Geschichtsdidaktik im Oktober 1991. *Geschichte in Wissenschaft und Unterricht*, 45(3), 164–176.
- Ji, Q. & Raney, A. A. (2016). Examining suspension of disbelief, perceived realism, and involvement in the enjoyment of documentary-style fictional films. *Projections*, 10(2), 125–142. <https://doi.org/10.3167/pr0j.2016.100207>
- De Jong, S. (2018). *The witness as object – video testimony in memorial museums*. New York: Berghahn Books.
- Kirchberg, V. & Tröndle, M. (2012). Experiencing exhibitions: A review of studies on visitor experiences in museums. *Curator: The Museum Journal*, 55(4), 435–452.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. „Aufgabenkultur“ und Unterrichtsgestaltung. In E. Klieme & J. Baumert (Hrsg.), *TIMSS-Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.
- Klimmt, C. & Vorderer, P. (2003). Media psychology „is not yet there“: Introducing theories on media entertainment to the presence debate. *Presence: Teleoperators and Virtual Environments*, 12(4), 346–359. <https://doi.org/10.1162/105474603322391596>
- Kohler, B. & Wacker, A. (2013). Das Angebots-Nutzungs-Modell: Überlegungen zu Chancen und Grenzen des derzeit prominentesten Wirkmodells der Schul- und Unterrichtsforschung. *Die deutsche Schule*, 105(3), 242–258.
- Kohlmeier, J. (2006). „Couldn’t she just leave?“. The relationship between consistently using class discussions and the development of historical empathy in a 9th grade world history course. *Theory & Research in Social Education*, 34, 34–57.
- Konijn, E. A. & Hoorn, J. F. (2005). Some like it bad. Testing a model for perceiving and experiencing fictional characters. *Media Psychology*, 7, 107–144.
- Körber Stiftung (2017). *Deutsche wollen aus Geschichte lernen*. Verfügbar unter: [https://lisa.gerda-henkel-stiftung.de/binaries/content/8511/2017-09-28\\_forsa-umfrage\\_geschichtsunterricht.pdf?t=1507558661](https://lisa.gerda-henkel-stiftung.de/binaries/content/8511/2017-09-28_forsa-umfrage_geschichtsunterricht.pdf?t=1507558661)
- Köster, M., Thünemann, H. & Zülsdorf-Kersting, M. (2017). Geschichtsunterricht als soziales System zwischen Angebot und Nutzung. *Zeitschrift für Geschichtsdidaktik*, 16(1), 218–234.

- Krämer, S. (2012). Zum Paradoxon von Zeugenschaft im Spannungsfeld von Personalität und Depersonalisierung. In M. Rössner & H. Uhl (Hrsg.), *Renaissance der Authentizität? Über die neue Sehnsucht nach dem Ursprünglichen* (S. 15–26). Bielefeld: transcript.
- Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409. [https://doi.org/10.1016/S0959-4752\(01\)00011-1](https://doi.org/10.1016/S0959-4752(01)00011-1)
- Krüger, T. (2020). Beutelsbach 2.0 - zehn Thesen zur politischen Bildung. In G. Hentges (Hrsg.), *Krise der Demokratie - Demokratie in der Krise? Gesellschaftsdiagnosen und Herausforderungen für die politische Bildung* (S. 177–193). Schwalbach/Ts.: Wochenschau.
- Kuchler, C. & Sommer, A. (2018). *Wirksamer Geschichtsunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Stuttgart: UTB.
- Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S. & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18(5), 468–482. <https://doi.org/http://dx.doi.org/10.1016/j.learninstruc.2008.06.008>
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 83–113). Münster: Waxmann.
- Landesinstitut für Schule und Medien Berlin-Brandenburg (2015). *Rahmenlehrplan für die Sekundarstufe I. Jahrgangsstufen 7-10. Geschichte*. Ministerium für Bildung, Jugend und Sport des Landes Brandenburg.
- Lanman, B. A. (1987). Oral history as an educational tool for teaching immigration and black history in American high schools. Findings and queries. *International Journal of Oral History*, 8, 122–135.
- Lee, P. (2004). „Walking backwards into tomorrow.“ Historical consciousness and understanding history. *International Journal of Historical Learning, Teaching and Research*, 4(1), 69–114.
- Lee, P. & Ashby, R. (2000). Progression in historical understanding among students Ages 7-14. In P. Stearns, P. Seixas & S. Wineburg (Hrsg.), *Knowing, teaching and learning history. National and international perspectives* (S. 199–222). New York University Press.

- Lee, P., Ashby, R. & Dickinson, A. (1996). Progression in children's ideas about history. In M. Hughes (Hrsg.), *Progression in learning* (S. 50–81). London: Bera Dialogues.
- Lehmann, J., Goussios, C. & Seufert, T. (2016). Working memory capacity and disfluency effect: An aptitude-treatment-interaction study. *Metacognition and Learning*, 11(1), 89–105. <https://doi.org/10.1007/s11409-015-9149-z>
- Lehmann, J. & Seufert, T. (2020). The interaction between text modality and the learner's modality preference influences comprehension and cognitive load. *Frontiers in Psychology*, 10(2820), 1–11. <https://doi.org/doi:10.3389/fpsyg.2019.02820>
- Levy, S. A. & Sheppard, M. (2018). „Difficult knowledge“ and the Holocaust in history education. In S. A. Metzger & L. McArthur Harris (Hrsg.), *The Wiley International Handbook of History Teaching and Learning* (S. 365–387). New York: Wiley-Blackwell.
- Liberman, N., Trope, Y. & Stephan, E. (2007). Psychological distance. In A. Kruglanski & T. Higgins (Hrsg.), *Social psychology: Handbook of basic principles 2* (2. Auflage, S. 353–383). New York: The Guilford Press.
- Llewellyn, K. R. & Ng-A-Fook, N. (2017). *Oral History and education*. Basingstoke: Palgrave Macmillan.
- Loderer, K., Pekrun, R. & Plass, J. L. (2020). Emotional foundations of game-based learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Hrsg.), *Handbook of game-based learning* (S. 111–151). Boston: MIT Press.
- Ma, W., Adesope, O. O., Nesbit, J. C. & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918. <https://doi.org/10.1037/a0037123>
- Maag Merki, K. (2021). Empirische Bildungsforschung im deutschsprachigen Raum. Rückblick und Ausblick. *Swiss Journal of Educational Research*, 43(1), 41–50. <https://doi.org/10.24452/sjer.43.1.4>
- Maggioni, L., Alexander, P. & VanSledright, B. (2004). At a crossroads? The development of epistemological beliefs and historical thinking. *European Journal of School Psychology*, 2(1), 169–197.
- Maggioni, L., Vansledright, B. & Alexander, P. (2009). Walking on the borders: A measure of epistemic cognition in history. *Journal of Experimental Education*, 77(3), 187–214. <https://doi.org/10.3200/JEXE.77.3.187-214>
- Memminger, J. (2022). Begegnungen mit Lebensgeschichten im Spannungsfeld von Empathie und kritischer Reflexion. *Geschichte lernen*, 25, 1–12.

- Meyer-Hamme, J. (2013). „I never liked history at school.“ Identitäten und Emotionen beim historischen Lernen. In J. Brauer & M. Lücke (Hrsg.), *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven* (S. 125–139). Göttingen: V&R unipress.
- Meyer-Hamme, J., Thünemann, H. & Zülsdorf-Kersting, M. (Hrsg.) (2016). *Was heißt guter Geschichtsunterricht?*. Schwalbach/Ts.: Wochenschau.
- Meyer, L. (2021). Widerstandsgeschichte auf Instagram: Was leistet das Projekt @ichbinsophiescholl? *Bundeszentrale für politische Bildung*. Verfügbar unter: <https://www.bpb.de/lernen/digitale-bildung/werkstatt/335601/widerstandsgeschichte-auf-instagram-was-leistet-das-projekt-ichbinsophiescholl/>
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2016). *Bildungsplan Baden-Württemberg. Fach Geschichte am Gymnasium*. Stuttgart. Verfügbar unter: [http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsw/export-pdf/depot-pdf/ALLG/BP2016BW\\_ALLG\\_GYM\\_G.pdf](http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_G.pdf)
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2014). *Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen. Geschichte*. Verfügbar unter: [https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp\\_SII/ge/KLP\\_GOSt\\_Geschichte.pdf](https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/ge/KLP_GOSt_Geschichte.pdf)
- Mrkva, K., Travers, M. & Van Boven, L. (2018). Simulational fluency reduces feelings of psychological distance. *Journal of Experimental Psychology*, 147(3), 354–376. <https://doi.org/10.1037/xge0000408>
- Nachtigall, V. & Rummel, N. (2021). Investigating students' perceived authenticity of learning activities in an out-of-school lab for social sciences: a replication study. *Instructional Science*, 49(6), 779–810. <https://doi.org/10.1007/s11251-021-09556-3>
- National Center for History in the Schools. (1996). National standards for history. Basic edition. University of California. Verfügbar unter: <https://phi.history.ucla.edu/nchs/history-standards/>
- Obens, K. & Geißler-Jagodzinski, C. (2010). „Dann sind wir ja auch die letzte Generation, die davon profitieren kann“. *Erste Ergebnisse einer empirischen Mikrostudie zur Rezeption von Zeitzeugengesprächen bei Jugendlichen/jungen Erwachsenen*. Verfügbar unter: [https://www.vielfalt-mediathek.de/wp-content/uploads/2020/12/forschungsbericht\\_zeitzeugengesprache.pdf](https://www.vielfalt-mediathek.de/wp-content/uploads/2020/12/forschungsbericht_zeitzeugengesprache.pdf)
- OECD (2005). School factors related to quality and equity - results from PISA 2000. Bonn.

- <https://doi.org/10.1787/9789264008199-en>
- Oral History Association (2013). *Principles and best practices for oral history education (4-12)*. Verfügbar unter: [http://www.oralhistory.org/wp-content/uploads/2014/04/2013-1411\\_Oral\\_History\\_ClassroomGuide\\_Update\\_V2.pdf](http://www.oralhistory.org/wp-content/uploads/2014/04/2013-1411_Oral_History_ClassroomGuide_Update_V2.pdf)
- Ostermann, E. (2002). *Die Authentizität des Ästhetischen. Studien zur ästhetischen Transformation der Rhetorik*. München: Wilhelm Fink.
- Paas, F., van Merriënboer, J. J. G. & van Gog, T. A. J. M. (2012). Designing instruction for the contemporary learning landscape. *APA Educational Psychology Handbook*, 3, 335–357. <https://doi.org/10.1037/13275-013>
- Pandel, H.-J. (1987). Dimensionen des Geschichtsbewußtseins. Ein Versuch, seine Struktur für Empirie und Pragmatik diskutierbar zu machen. *Geschichtsdidaktik*, 12, 130–142.
- Pandel, H.-J. (2005). *Geschichtsunterricht nach PISA. Kompetenzen, Bildungsstandards und Kerncurricula*. Schwalbach/Ts.: Wochenschau.
- Pandel, H.-J. (2009). Geschichtsbewusstsein. In U. Mayer, H.-J. Pandel, G. Schneider & B. Schönemann (Hrsg.), *Wörterbuch Geschichtsdidaktik* (S. 80). Schwalbach/Ts.: Wochenschau.
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G. et al. (2017). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205–242. <https://doi.org/10.3102/0034654317743198>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315–341. <https://doi.org/https://doi.org/10.1007/s10648-006-9029-9>
- Pekrun, R. (2016). Academic emotions. In K. Wentzel & D. B. Miele (Hrsg.), *Handbook of Motivation at School* (S. 120–144). New York: Routledge.
- Pekrun, R. & Linnenbrink-Garcia, L. (2014). *International handbook of emotions in education*. New York: Routledge.
- von Plato, A. (2001). Chancen und Gefahren des Einsatzes von Zeitzeugen im Unterricht. *BIOS Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen*, 14(2), 134–139.
- Potter, D. & Morris, D. S. (2017). Family and schooling experiences in racial/ethnic academic achievement gaps: A cumulative perspective. *Sociological Perspectives*, 60(1), 132–167. <https://doi.org/10.1177/0731121416629989>
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring,

- A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum - zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446.  
<https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Rogh, W. & Kleickmann, T. (2020). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft*, 48(3), 303–318.  
<https://doi.org/10.1007/s42010-020-00072-w>
- Renkl, A. (2015). Wissenserwerb. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 3–24). Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-41291-2>
- Richardson, A. (2021). Touching distance: Young people’s reflections on hearing testimony from Holocaust survivors. *Journal of Modern Jewish Studies*, 1–24.  
<https://doi.org/10.1080/14725886.2021.1874692>
- Riegel, U. & Kindermann, K. (2016). Why leave the classroom? How field trips to the church affect cognitive learning outcomes. *Learning and Instruction*, 41, 106–114.  
<https://doi.org/10.1016/j.learninstruc.2015.10.004>
- Rodenhäuser, L. (2012). *Zwischen Affirmation und Reflexion. Eine Studie zur Rezeption von Zeitzeugen in Geschichtsdokumentationen*. Berlin: Lit.
- Roscigno, V. J. & Ainsworth-Darnell, J. W. (1999). Race, cultural capital, and educational resources: Persistent inequalities and achievement returns. *Sociology of Education*, 72(3), 158–178. <https://doi.org/10.2307/2673227>
- Rosenzweig, R. & Thelen, D. P. (1998). *The presence of the past: Popular uses of history in American life*. New York: Columbia University Press.
- Rüsen, J. (1994). *Historisches Lernen*. Köln: Böhlau.
- Rüsen, J. (2005). *History: Narration, interpretation, orientation*. New York: Berghahn Books.
- Rüsen, J. (2013). Die Macht der Gefühle im Sinn der Geschichte. Theoretische Grundlagen und das Beispiel des Trauerns. In J. Brauer & M. Lücke (Hrsg.), *Emotionen, Geschichte und historisches Lernen - Geschichtsdidaktische und geschichtskulturelle Perspektiven* (S. 27–47). Göttingen: V&R unipress.
- Rüsen, J. (2020). *Geschichte denken. Erläuterungen zur Historik*. Wiesbaden: Springer.
- Sabrow, M. (2007). Veränderte Verhältnisse - zum historischen Wandel der Auseinandersetzung mit dem Nationalsozialismus in Deutschland. *Der Umgang mit der Zeit des Nationalsozialismus. Perspektiven des Erinnerns* (S. 63–81). Landeshauptstadt München.

- Sabrow, M. (2012). Der Zeitzeuge als Wanderer zwischen den Welten. In M. Sabrow & N. Frei (Hrsg.), *Die Geburt des Zeitzeugen nach 1945* (S. 13–32). Göttingen: Wallstein.
- Sakaki, M., Ueno, T., Ponzio, A., Harley, C. & Mather, M. (2019). Emotional arousal amplifies competitions across goal-relevant representation: A neurocomputational framework. *Cognition*, *187*, 108–125.
- Sauer, M. (2006). *Geschichte unterrichten. Eine Einführung in die Didaktik und Methodik*. Seelze: Kallmeyer.
- Saupe, A. (2017). Historische Authentizität: Individuen und Gesellschaften auf der Suche nach dem Selbst. *H-Soz.-Kult*, *15*, 50–95.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*, 299–323.
- Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The role of interest in learning and development*. (S. 183–212). Hillsdale: Lawrence Erlbaum Associates.
- Schneider, M. & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology*, *49*(9), 1639–1649.  
<https://doi.org/10.1037/a0030976>
- Schreiber, W. (2018). Herzstück des Geschichtsunterrichts ist historisch Denken lernen. In C. Kuchler & A. Sommer (Hrsg.), *Wirksamer Geschichtsunterricht* (S. 149–161). Baltmannsweiler: Schneider Verlag Hohengehren.
- Schreiber, W. & Árkossy, K. (2009). *Zeitzeugengespräche führen und auswerten. Historische Kompetenzen schulen*. Neuried: ars una.
- Schreiber, W., Körber, A., von Borries, B., Krammer, R., Leutner-Ramme, S., Mebus, S. et al. (2006). *Historisches Denken. Ein Kompetenz-Strukturmodell*. Neuried: ars una.
- Schwan, S. & Dutz, S. (2020). How do visitors perceive the role of authentic objects in museums? *Curator*, *63*(2), 217–237. <https://doi.org/10.1111/cura.12365>
- Seidel, T. (2008). Schuleffektivitätskriterien in der internationalen empirischen Forschung. *Zeitschrift für Erziehungswissenschaften*, *12*(3), 349–367.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie: Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, *60*(6), 850–866.
- Seidel, T. & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (S. 253–296). Weinheim: Beltz.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The



- role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Seixas, P. (2017). A model of historical thinking. *Educational Philosophy and Theory*, 49(6), 593–605.
- Seixas, P. & Morton, T. (2013). *The big six historical thinking concepts*. Toronto: Nelson Education.
- Shaffer, D. W. & Resnick, M. (1999). „Thick“ authenticity: New media and authentic learning. *Journal of Interactive Learning Research*, 10(2), 195–215. <https://doi.org/10.1016/j.stomax.2007.01.005>
- Sheppard, M., Katz, D. & Grosland, T. (2015). Conceptualizing emotions in social studies education. *Theory & Research in Social Education*, 43(2), 147–178. <https://doi.org/10.1080/00933104.2015.1034391>
- Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L. & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, 57(1), 31–54. <https://doi.org/10.1080/00461520.2021.1939700>
- Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology*, 59(2), 205–216. <https://doi.org/10.1037/0022-006X.59.2.205>
- Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, 27(1), 5–32.
- Snow, R. E., Kanfer, R. & Cudeck, R. (1989). Cognitive-conative aptitude interactions in learning. In R. Kanfer, P. L. Ackerman & R. Cudeck (Hrsg.), *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences*. (S. 435–475). Hillsdale, NJ: Lawrence Erlbaume Associates.
- Sochatzy, F. (2010). Zeitzeugen im Geschichtsunterricht. In M. Ventzke, S. Mebus, W. Schreiber, B. von Borries, A. Körber, C. Pfluger et al. (Hrsg.), *Geschichte denken statt pauken in der Sekundarstufe II. 20 Jahre nach der friedlichen Revolution: Deutsche und europäische Perspektiven im gymnasialen Geschichtsunterricht*. Radebeul: Sächsisches Bildungsinstitut.
- Van Straaten, D., Wilschut, A. & Oostdam, R. (2016). Making history relevant to students by connecting past, present and future: A framework for research. *Journal of Curriculum Studies*, 48(4), 1–24. Routledge. <https://doi.org/10.1080/00220272.2015.1089938>
- Van Straaten, D., Wilschut, A., Oostdam, R. & Fukkink, R. (2019). Fostering students' appraisals of the relevance of history by comparing analogous cases of an enduring

- human issue: A quasi-experimental study. *Cognition and Instruction*, 37(4), 512–533. Routledge. <https://doi.org/10.1080/07370008.2019.1614590>
- Südwestrundfunk (SWR). (2021). @ichbinsophiescholl. Verfügbar unter: <https://www.swr.de/unternehmen/ich-bin-sophie-scholl-instagram-serie-102.html>
- Szucs, D. & Goswami, U. (2007). Educational neuroscience: Defining a new discipline for the study of mental representations. *Mind, Brain and Education*, 1(3), 114–127.
- Tetzlaff, L., Schmiedek, F. & Brod, G. (2021). Developing personalized education: A dynamic framework. *Educational Psychology Review*, 33(3), 863–882.
- Thünemann, H. (2021). Erkenntnisse zu Unterrichtsmethoden im Geschichtsunterricht. In G. Weißeno & B. Ziegler (Hrsg.), *Handbuch Geschichts- und Politikdidaktik* (S. 1–14). Wiesbaden: Springer.
- Trautwein, U., Bertram, C., Brauch, N., Hirsch, M., Klausmeier, K., Merkt, M. et al. (2017). *Kompetenzen historischen Denkens erfassen*. Münster: Waxmann.
- Trautwein, U., Schreiber, W. & Hasenbein, L. (2021). Geschichtsdidaktische Forschung und allgemeine Unterrichtsforschung: Unbekannte Verwandte? *Unterrichtswissenschaft*, 49, 261–272. <https://doi.org/10.1007/s42010-021-00120-z>
- Trofanenko, B. (2017). „We tell stories“: Oral history as a pedagogical encounter. In K.R. Llewellyn & N. Ng-A-Fook (Hrsg.), *Oral history and education* (S. 149–165). New York: Palgrave Macmillan. <https://doi.org/10.2307/3119791>
- Trope, Y. & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463. <https://doi.org/10.1037/a0018963>
- Trškan, D. (2016). *Oral history education—dialogue with the past*. Ljubljana: Slovenian National Commission for UNESCO.
- Ueding, G. (2012). Authentizität. *Historisches Wörterbuch der Rhetorik* (Band 10, S. 81–87). Berlin: De Gruyter.
- Vieluf, S., Praetorius, A.-K., Rakoczy, K., Kleinknecht, M. & Pietsch, M. (2020). Angebots-Nutzungs-Modelle als Rahmung des Unterrichts. *Zeitschrift für Pädagogik*, 66(1), 63–80.
- Vorderer, P. (1992). *Fernsehen als Handlung. Fernsehfilmrezeption aus motivationspsychologischer Perspektive*. Berlin: Edition Sigma.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4, 45–69.
- Vosniadou, S. & Mason, L. (2012). Conceptual change induced by instruction: A complex interplay of multiple factors. In K. R. Harris, S. Graham, T. Urdan, J. Royer & M.

- Zeidner (Hrsg.), *APA Educational Psychology Handbook, Vol. 2: Individual differences and cultural and contextual factors* (S. 221–246). American Psychological Association.
- Wacker, A., Fend, H., Bohl, T., Rehm, M., Engel, J. & Rohlfs, C. (2020). Die Einheit der Bildungswissenschaften und die Vielfalt ihrer Disziplinen und Forschungsfelder. *heiEDUCATION*, 6, 11–43.
- Wagner, W., Göllner, R., Helmke, A., Trautwein, U. & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, 28, 1–11. <https://doi.org/10.1016/j.learninstruc.2013.03.003>
- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B. & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5), 705–721. <https://doi.org/10.1037/edu0000075>
- Weber, P. (2021). Violations of realism in experiencing stories. Salience, tolerance and role in narrative persuasion. Verfügbar unter: [https://www.researchgate.net/profile/Patrick\\_Weber18](https://www.researchgate.net/profile/Patrick_Weber18)
- Wehling, H.-G. (1977). Konsens à la Beutelsbach? Nachlese zu einem Expertengespräch. In S. Schiele & H. Schneider (Hrsg.), *Das Konsensproblem in der politischen Bildung* (S. 173–184). Stuttgart: Klett.
- Weinert, F. (2001). *Leistungsmessung in Schulen*. Berlin: Beltz.
- Westdeutscher Rundfunk (WDR) (2021). Die WDR History App holt die Geschichte ins Klassenzimmer. *WDR AR 1933-1945*. Verfügbar unter: <https://www1.wdr.de/fernsehen/unterwegs-im-westen/ar-app/index.html>
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461–481. <https://doi.org/10.1037/0033-2909.91.3.461>
- Whitman, G. (2004). *Dialogue with the past*. Walnut Creek: Altamira.
- Wierling, D. (1997). Oral History. In K. Bergmann, K. Fröhlich, A. Kuhn, J. Rösen & G. Schneider (Hrsg.), *Handbuch der Geschichtsdidaktik* (S. 236–239). Seelze-Velber: Kallmeyer.
- Wilhelm, M., Rehm, M. & Reinhardt, V. (2018). Qualitätvoller Fachunterricht. In C. Kuchler & A. Sommer (Hrsg.), *Wirksamer Geschichtsunterricht* (S. 11–19). Baltmannsweiler: Schneider Verlag Hohengehren.
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in

the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73–87.

Wolter, H. (2013). *Wir alle sind Geschichte: Lehrerhandreichung zu Oral History im Geschichtsunterricht*. Universität Regensburg.

Zabori, T. (2021). Instagram-Projekt @ichbinsophiescholl - Einsatz im Unterricht. *planet schule*. Verfügbar unter: [https://www.planet-schule.de/sf/spezial/spezial\\_ichbinsophiescholl\\_unterricht.php](https://www.planet-schule.de/sf/spezial/spezial_ichbinsophiescholl_unterricht.php)

Zülsdorf-Kersting, M. (2020). Quality of history education and its criteria – between generic and disciplinary perspectives. *Unterrichtswissenschaft*, 48(3), 385–407.  
<https://doi.org/10.1007/s42010-020-00073-9>

