



Lautes Denken

Julia Thyroff

Inhalt

1	Definitionen des Gegenstands	2
2	Hintergründe und Leistungsfähigkeit der Methode	2
3	Varianten von Anliegen und Erhebungssettings	4
4	Lautes Denken in der geschichts- und politikdidaktischen Forschung	5
5	Reichweiten und Limitationen	10
	Literatur	12

Zusammenfassung

Im Beitrag wird Lautes Denken als Methode vorgestellt und in seiner Bedeutung für die geschichts- und politikdidaktische Forschung verortet. Hierfür werden zunächst fachunspezifisch Hintergründe, Charakteristika und Leistungsfähigkeit der Methode eingeordnet. Im Anschluss werden bisherige Einsatzfelder Lauten Denkens in der Geschichts- und Politikdidaktik dargestellt. Während die Methode in der Geschichtsdidaktik in zahlreichen Zusammenhängen zum Einsatz kommt, ist sie in der Politikdidaktik noch weniger gebräuchlich. Die Zusammenschau mündet schliesslich in eine Diskussion der Reichweiten und Limitationen der Methode sowie die Benennung von Entwicklungsfeldern für beide Disziplinen.

Schlüsselwörter

Prozessbegleitendes Lautes Denken · Nachträgliches Lautes Denken · Stimulated Recall · Erhebungsmethode · Denkprozesse · Historisches Denken · Planungsverhalten

J. Thyroff (✉)

Zentrum Politische Bildung und Geschichtsdidaktik, PH FHNW, Aarau, Schweiz

E-Mail: julia.thyroff@fhnw.ch

1 Definitionen des Gegenstands

Lautes Denken ist eine Erhebungsmethode, die in zahlreichen Disziplinen eingesetzt wird. Im Kern besteht die Methode darin, dass beforschte Personen aussprechen, was ihnen beim Erleben einer Situation oder beim Bewältigen einer Aufgabe durch den Kopf geht (Someren et al. 1994, S. 42; Weidle und Wagner 1982, S. 81). Lautes Denken soll „Einblicke in die Gedanken, Gefühle und Absichten einer lernenden und/oder denkenden Person“ ermöglichen (Konrad 2010, S. 475). Anfänglich diente Lautes Denken der Beobachtung kognitiver Vorgänge in der Psychologie und erlebte dann eine Renaissance mit der Problemlöseforschung, wo es Denkprozesse und Lösungsstrategien sichtbar machen sollte (Konrad 2010, S. 477–478; Someren et al. 1994, S. 8–9, 13–15, 29–32). Darüber hinaus eignet sich die Methode für Forschungsvorhaben unterschiedlicher Art, wann immer es darum geht, „to validate or construct theories of cognitive processes“ (Someren et al. 1994, S. 9, 13–14). Zu den Einsatzfeldern der Methode gehören die Entscheidungsforschung, Spracherwerbs- und Leseforschung, Unterrichtsforschung u. v. m. und sie lässt sich sowohl mit qualitativen wie auch quantitativen Auswertungsstrategien kombinieren (Konrad 2010, S. 480–483; Stark 2010; Weidle und Wagner 1982, S. 81–82, 91–98).

Beim Lauten Denken handelt es sich streng genommen nicht um eine einheitliche Methode, sondern um den Überbegriff für eine Vielzahl von Einsatzvarianten (Stark 2010, S. 58). Lautes Denken existiert in zwei Spielarten, als prozessbegleitendes und als nachträgliches Lautes Denken, letzteres teilweise auch als *Stimulated Recall* bezeichnet und durch Impulsmedien unterstützt (Ericsson und Simon 1980, S. 218, 1984, S. 16; Konrad 2010, S. 476–477, 481; Lyle 2003; Wallach und Wolf 2001; Weidle und Wagner 1982). Weiter lassen sich Varianten dahingehend unterscheiden, ob die zu verbalisierende Information bereits in sprachlicher Form im Gedächtnis existiert oder ob „one or more mediating processes“ notwendig werden, weil bildliche Information in Sprache transformiert werden muss oder weil die Aufgabenstellung vorsieht, gezielt nur bestimmte Informationen zu verbalisieren (Ericsson und Simon 1980, S. 219–220). (für dieses Kapitel vgl. auch Thyroff 2020, S. 195–201).

2 Hintergründe und Leistungsfähigkeit der Methode

Um die Beschreibung Lauten Denkens haben sich Ericsson und Simon verdient gemacht (1980, 1984, 1993). Sie und andere gehen von einer Dreiteilung des menschlichen Gedächtnisses in „*sensory stores*“, „*short-term memory*“ und „*long-term memory*“ aus, wobei Lautes Denken Einblicke in das Kurzzeitgedächtnis gewähre (z. B. Ericsson und Simon 1993, S. 11; Bannert 2007, S. 134–135; Konrad 2010, S. 478; Wallach und Wolf 2001, S. 13). Inwiefern Lautes Denken unmittelbare Rückschlüsse auf Denkprozesse erlaubt, wird jedoch kontrovers eingeschätzt. Wallach und Wolf unterscheiden zwei Arten von Einwänden: solche zur „Non-Veridikalität“, die „das nicht wahrheitsgetreue Berichten über den Denkprozeß“ beanstanden, und solche zur „Reaktivität“ Lauten Denkens, die beanstanden, „daß der Prozeß,

während dessen verbalisiert wird, sich durch die Verbalisierung ändert“ (Wallach und Wolf 2001, S. 17).

Im Bereich der Reaktivität finden sich Hinweise, dass Lautes Denken zu einer langsameren Bearbeitung der gestellten Aufgabe führt (Deffner 1984, S. 143; Someren et al. 1994, S. 33) und den Verlauf der Aufgabenbearbeitung beeinflussen kann (Wallach und Wolf 2001, S. 18–19; Weidle und Wagner 1982, S. 99–100). Proband/-innen, nach dem Einfluss der Methode befragt, berichten wechselweise keine Effekte (Litten 2017, S. 204) oder aber eine erhöhte Sorgfalt, Strukturiertheit und Konzentration bei der Aufgabenbearbeitung (Deffner 1984, S. 20–21) – dennoch sei die Qualität der Ergebnisse nicht vom Lauten Denken beeinflusst worden (S. 143–144). Einschätzungen der Methode fallen also widersprüchlich aus, in der Gesamtschau bestehen aber „ernst zu nehmende empirische Hinweise auf eine Reaktivität Lauten Denkens“ (Wallach und Wolf 2001, S. 18).

Im Bereich der Non-Veridikalität liegen Erkenntnisse vor, dass die Verbalisierung sprachlicher Informationen besser gelingt als bildlicher (Someren et al. 1994, S. 122; Stark 2010, S. 73–74). Auch bei zeitgleichem Lesen ist das Laute Denken eingeschränkt und kann dazu führen, dass Proband/-innen „simultane[.] Nacherzählungen“ der gelesenen Texte produzieren, statt im eigentlichen Sinne Gedanken zu äussern (Stark 2010, S. 63, 74). Weiter gibt es Hinweise, dass die Verbalisierung deklarativen Wissens besser gelingt als prozeduralen Wissens (Wallach und Wolf 2001, S. 11). Zudem droht eine unvollständige Verbalisierung, sofern kognitive Prozesse sehr komplex sind oder sehr schnell vonstattengehen (Someren et al. 1994, S. 33, 122; Weidle und Wagner 1982, S. 84). „Es ist nicht möglich, alles, was einem durch den Kopf geht, auch auszusprechen. Viele Gedanken tauchen nur sehr kurz auf und verschwinden wieder, noch bevor man sie in Worte fassen kann. [...] Das bedeutet, dass Lautdenkprotokolle naturgemäß immer unvollständig und fragmentarisch sind“ (Stark 2010, S. 73).

Überdies gelingt das Laute Denken unterschiedlich gut je nach Person (Reich 2009, S. 335; Saaman 2001, S. 125, 128–129, 135; Someren et al. 1994, S. 35–36) oder zugrunde liegendem Aufgabentyp (Wallach und Wolf 2001, S. 19). Speziell beim nachträglichen Lauten Denken zeigt sich zudem, dass Proband/-innen nicht zwingend über Strategien Auskunft geben (können), die sie zuvor eingesetzt haben (Rogiers et al. 2020, S. 315). Zuletzt lässt sich festhalten, dass Lautes Denken nur solche Inhalte des Gedächtnisses zutage fördern kann, die den Proband/-innen bewusst sind (Konrad 2010, S. 486; Wallach und Wolf 2001, S. 14; Weidle und Wagner 1982, S. 83). Gerade Prozesse, die stark automatisiert ablaufen, entziehen sich damit einer Verbalisierung (Stark 2010, S. 58).

Es können also eine Reihe von Einwänden gegen die Methode vorgebracht werden. Gerade quantitativ Forschende bemängeln, dass die Methode „oft nicht vollständig kontrollierbar, nur bedingt reproduzierbar und vor allem wenig valide sei“ (Konrad 2010, S. 480). Indes gelten etliche Einwände, die gegen Lautes Denken vorgebracht werden, gleichermassen für weitere Erhebungsmethoden, die auf Verbalisierungen setzen, weshalb es weniger darum gehen sollte, „die genannten Einschränkungen auszuschalten, sondern vielmehr bewusst zu machen und zu reflektieren“ (Thyroff 2020, S. 207) – ebenso wie die Chancen der Methode. Hierzu zählt

ihre „ausgeprägte Prozessbezogenheit“, die es erlaubt, unmittelbare Einblicke in kognitive Vorgänge zu erhalten, „die anders nicht erhoben werden könnten“ (Konrad 2010, S. 481, 485). Eng damit verbunden ist das Potenzial, Proband/-innen unmittelbar in Interaktion mit Materialien oder ihrer physischen Umwelt zu beobachten (Thyroff 2020, S. 214–216, 617–619; Wise 2011, S. 10, 21). Auch erlaubt Lautes Denken den Proband/-innen, in relativer Ungestörtheit von den Forschenden ihre Gedanken zu verbalisieren (Someren et al. 1994, S. 25), also „*to explore and in essence to create their own prompts*“ (Wise 2011, S. 20). Mehrere Autor/-innen betonen zudem, dass Lautes Denken ausgesprochen reichhaltige Daten erzeuge: „Streng genommen ist die Bezeichnung ‚Lautes Denken‘ nicht vollständig zutreffend, da neben Produkten des Denkens auch Wahrnehmungen, Gefühle, Empfindungen und Gedanken über das eigene Denken und Handeln (Metakognitionen) verbalisiert werden“ (Wallach und Wolf 2001, S. 9; dazu auch Lange 2011, S. 113; Weidle und Wagner 1982, S. 82). (für dieses Kapitel vgl. auch Thyroff 2020, S. 201–209)

3 Varianten von Anliegen und Erhebungssettings

Durch gezielte Erhebungssettings lassen sich die Chancen und Herausforderungen der Methode unterstützen oder reduzieren – je nach Anliegen. So wird das Auftreten von metakognitiven Äusserungen in manchen Studien geschätzt (z. B. Kaiser 2018; Saaman 2001, S. 125, 128–129, 135), andernorts wird es zu unterbinden versucht, etwa mit dem Auftrag, zwar laut zu denken, nicht aber Verhalten und Gedanken zu interpretieren (Someren et al. 1994, S. 25, 44). Mitunter gelten reflektierende Äusserungen „im Sinne eines ‚Theoretisierens‘ über sich selbst“ gar nicht als Lautes Denken im eigentlichen Sinne (Weidle und Wagner 1982, S. 83; dazu auch Stark 2010, S. 60–61). Gerade in Bezug auf nachträgliches Lautes Denken bemängelt Lyle, dass nachträgliches Filtern („*measure of ‚sanitising‘*“), Ordnen und Reflektieren („*intermediate ordering of reflections on reasoning*“) auftreten könne (Lyle 2003, S. 864–866). Lyle empfiehlt, solche Effekte durch das Erhebungsdesign zu minimieren: „*reduce anxiety; limit the perception of judgemental probing; reduce the intrusion into the action; stimulate rather than present a novel perspective/insight; make the retrospection as immediate as possible; allow the subject a relatively unstructured response; and employ an ‚indirect‘ route to the focus of the research*“ (Lyle 2003, S. 865–866).

Andere Autor/-innen streben hingegen metakognitive Äusserungen dezidiert an. Insbesondere unter dem Sammelbegriff *Stimulated Recall* zielen Autor/-innen häufig weniger auf die Rekonstruktion von Gedanken als auf Reflexionen und Beurteilungen des Gesehenen, beispielsweise verstanden als „Selbstbeobachtung professionellen Handelns“ (Maag Merki und Werner 2014, S. 755) oder „laute[s] Denken über eigene Handlungspraxis“ (Riegel 2013, S. 12).

Essenziell beim Einsatz Lauten Denkens ist die erteilte Instruktion. Während in manchen Studien einschränkende Hinweise gegeben werden, etwa, um das Auftreten von Metakognitionen zu unterbinden, werden andernorts möglichst breite Aufforderungen erteilt, z. B. „*to verbalize everything that they ‚heard themselves*

thinking“ (Wineburg 1998, S. 323, ähnlich breit 1991, S. 75). Mitunter wird ergänzend betont, dass wirklich alle Gedanken interessant seien, um zu vermeiden, dass Teilnehmende trotz breiter Aufforderung in Antizipation des Forschungsinteresses nur selektiv verbalisieren (Stark 2010, S. 69; Thyroff 2020, S. 221).

In vielen Studien findet zudem eine Einübung Lauten Denkens statt, um Proband/-innen mit der Methode vertraut zu machen und ihnen Sicherheit darüber zu geben, dass sie diese ‚richtig‘ anwenden (Fromann 2005, S. 4; Schnell 2016, S. 34; Someren et al. 1994, S. 35–36, 43; Stark 2010, S. 68–69; Wineburg 1991, S. 75). Auch werden bisweilen unterstützende „*guidelines*“ abgegeben, z. B. „*say whatever’s on your mind*“, „*don’t overexplain or justify*“, „*don’t worry about complete sentences*“ (Wineburg 1998, S. 323), oder Beispielformulierungen angeboten, z. B. „jetzt überlege ich gerade wie“ (Fromann 2005, S. 4). Vielerorts erweist es sich zudem als notwendig, Proband/-innen während der Erhebung an das Laute Denken zu erinnern (Litten 2017, S. 182, 204; Lange 2011, S. 114; Wineburg 1991, S. 75; Stark 2010, S. 70; Schnell 2016, S. 34), während andernorts Lautes Denken ohne Anwesenheit von Forschenden selbsttätig abläuft (Thyroff 2020, S. 222, 273).

Zu klären ist auch die Frage, ob Proband/-innen im Falle einer Konfrontation mit Texten nicht nur laut denken, sondern auch laut lesen sollen (so z. B. bei Thyroff 2020; Wineburg 1991, S. 75, 1998, S. 323). Lautes Lesen macht beobachtbar, wann überhaupt gelesen wird und in welchen Momenten der Lektüre welche Gedanken entstehen (Thyroff 2020, S. 222), beeinflusst aber in komplexer Weise den Rezeptionsprozess. So kann es die Lesedauer verlängern und zur Verfestigung von Falschannahmen führen, umgekehrt aber auch Aufmerksamkeit und Textwahrnehmung intensivieren (Stark 2010, S. 76–77).

Bestimmte Bedingungen können das parallele Laute Denken erschweren oder verunmöglichen. In solchen Fällen bieten sich das nachträgliche Laute Denken oder Mischformen an. Um eine Parallelität von Denken und Lesen zu vermeiden, sind Settings gebräuchlich, in welchen Proband/-innen Texte sequenzweise lesen und jeweils unmittelbar darauf laut denken sollen (Stark 2010, S. 63–64; Wineburg 1991, S. 75; Wolfe und Goldman 2005, S. 467). Nachträgliches Lautes Denken bietet sich vor allem bei Situationen mit einer hohen Interaktivität und Kommunikationsdichte an, die ein paralleles Lautdenken gänzlich unmöglich machen, so etwa der Unterricht (Lyle 2003, S. 862; Weidle und Wagner 1982, S. 81–82). Soll *Stimulated Recall* zum Einsatz kommen, stellt die Auswahl der Impulsmedien einen weiteren zu reflektierenden Faktor dar (Lyle 2003, S. 863).

4 Lautes Denken in der geschichts- und politikdidaktischen Forschung

In der Geschichtsdidaktik kam Lautes Denken bereits in etlichen Zusammenhängen zum Einsatz, während es in der Politikdidaktik bis anhin weniger gebräuchlich ist. In der Gesamtschau lassen sich bisherige Studien in sechs Bereiche gliedern:

- (1) Studien zu historischem und politischem Denken und Lernen von Schüler/-innen und Erwachsenen im Rahmen konstruierter Aufgabensettings,
- (2) Studien zur Validierung von Tests,
- (3) Studien zu Denkprozessen von Schüler/-innen im Geschichtsunterricht,
- (4) Studien zu Denkprozessen von Besuchenden an geschichtskulturellen Orten,
- (5) Studien zum Planungsverhalten von Geschichtslehrpersonen,
- (6) Studien zu subjektiven Konzepten und Theorien von (angehenden) Lehrpersonen.

Zu (1): Am häufigsten kam Lautes Denken bislang zum Einsatz, um im Rahmen konstruierter Aufgabensettings fachspezifische Denk- und Lernprozesse zu erhellen, vor allem in der Geschichtsdidaktik, wo es zur Erforschung historischen Denkens und seiner Teilaspekte gebraucht wird. Damit wird der Forderung Körbers entsprochen, dass Forschung zum historischen Denken unbedingt „Prozesse des historischen Denkens *in actu* in den Blick nehmen“ müsse, während in retrospektiv erhobenen ‚fertigen‘ Produkten Prozesse nicht mehr unmittelbar zugänglich seien (Körber 2006, S. 191, kursiv im Original).

Wegbereitend in diesem Zusammenhang sind die Studien Wineburgs (1991, 1998, 2001), in welchen Personen unterschiedlicher Professionalitätsgrade (Schüler/-innen, Historiker/-innen) in der Auseinandersetzung mit Materialien laut denken sollten, beispielsweise, um herauszufinden, wie sie „*construct an understanding of historical events from a group of fragmented and contradictory documents*“ (Wineburg 1991, S. 74–75). In der Folge Wineburgs finden sich seither zahlreiche Studien, in welchen das Denken von Proband/-innen im Umgang mit Materialien und bei der Erledigung von Aufträgen beobachtet wurde (z. B. Cho et al. 2018; Jay 2021; Shreiner 2019). Bereits mehrfach gerieten etwa Denkprozesse in Auseinandersetzung mit kontrastierenden Darstellungstexten in den Blick (z. B. Houwen et al. 2020; Iordanou et al. 2020). Beispielsweise wurden von Wolfe und Goldman (2005) Jugendliche mit zwei „*contradictory texts explaining the Fall of Rome*“ (S. 467) konfrontiert und sollten nach der Lektüre jeden Satzes laut denken.

In der deutschsprachigen Geschichtsdidaktik wurde der Umgang mit Materialien von Lange (2011) beforscht. Sie setzte prozessbegleitendes und nachträgliches Lautes Denken ein, ergänzt um Concept Mapping und Gruppendiskussionen, um „*individuelles und kollaboratives Lernen mit historischen Bildquellen als Produkt und Prozess*“ zu erhellen (S. 108; kursiv im Original). Und an der Schnittstelle geschichts- und politikdidaktischer Anliegen steht eine Untersuchung Shreiners (2014). Dort interessierte, „*how eight political scientists and eight high school students employed historical knowledge to reason about a political issue*“ (S. 313). Die Proband/-innen wurden mit Materialien und Urteilsfragen konfrontiert und sollten in fünf verschiedenen Phasen prozessbegleitend wie auch nachträglich laut denken (S. 313, 319–320, 328).

Bislang existierende Studien haben explorativen Charakter und arbeiten mit vergleichbar kleinen Fallzahlen. Der Einsatz Lauten Denkens ist aufwändig (Kaliski et al. 2015, S. 204), insbesondere wenn die Methode mit einer differenzierten qualitativen Auswertung verbunden wird. Umso sinnvoller und vielversprechender

ist unter diesen Umständen eine gezielte, kontrastive Samplingstrategie entlang des Experten-Laien-Paradigmas, wie sie in Tradition Wineburgs bereits in etlichen Projekten zum Einsatz kam (z. B. McGrew 2021; Shreiner 2014; Wise 2011; für politisches Denken: Bernstein 2010). Auf diese Weise wird es möglich, auch mit kleinen Fallzahlen das mögliche Spektrum fachlichen Denkens auszuloten und prozessbezogene Einsichten zu generieren, die sich mittels alternativer Erhebungsformen nicht erzielen liessen.

Zu (2): Ausprägungen fachspezifischen Denkens rückten auch in Studien in den Blick, die sich der validierenden Begleitforschung zu standardisierten Tests widmeten. Für die Geschichtsdidaktik reflektiert wird dieses Forschungsfeld bei Kaliski, Smith und Huff (2015). Herausforderungen bei der Testkonstruktion für Geschichte bestünden in der Komplexität des adressierten Konstrukts (z. B. *Historical Thinking*), seiner vielfältigen und unterschiedlich verstandenen Teildimensionen und seiner „*higher-order nature*“, die sich der Validierung durch „*traditional content validity procedures*“ entziehe. Neben einer Faktorenanalyse zur Bestimmung der von einem Test adressierten Dimensionen historischen Denkens könne Lautes Denken dienen, um „*a cognitive model of task performance*“ zu gewinnen (S. 196). Der Einsatz der Methode sei ressourcenintensiv und werde oft vermieden, liefere aber „*a strong evidentiary argument for the validity of history assessment test scores*“ (S. 204).

In diesem Sinn kam Lautes Denken bereits mehrfach zum Einsatz, um herauszufinden, inwieweit Fragen eines Tests die gewünschten historischen Denkprozesse hervorriefen – mit unterschiedlichen Ergebnissen (z. B. Ercikan et al. 2015; Smith et al. 2019). Etwa setzte Reich (2009) Lautes Denken ein, hier in Kombination mit nachträglichen Interviews, um herauszufinden, inwiefern Multiple-Choice-Fragen eines Geschichtstests „*knowledge and skills*“ hervorriefen (S. 325). Er kam zu dem Schluss, dass sich die der Testkonstruktion zugrunde liegenden „*second-order historical concepts*“ nicht in den Äusserungen der Teilnehmenden widerspiegelten (S. 347–348).

Ebenfalls zur Erfassung von Konzepten wurde Lautes Denken in der Politikdidaktik von Hartmann (2015) erwogen – dies zugleich der einzige bekannte Einsatz Lauten Denkens in der deutschsprachigen Politikdidaktik. Hartmann interessierte das konzeptuelle politische Wissen, das während einer Testbearbeitung bei Schüler/-innen aktiviert wird (S. 138). Im Gegensatz zu den genannten geschichtsdidaktischen Studien kam Hartmann zum Schluss, dass Lautes Denken für ihre Zwecke „nur eingeschränkt geeignet [sei] – insbesondere wegen des Verzichts der Methode auf vertiefendes inhaltliches Nachfragen und des fehlenden Gesprächscharakters. Zu beobachten waren vor allem die Tendenz der Befragten zu vereinfachten Wertungen („*Antwort X ist ja Quatsch.*“) und die Neigung, nicht so sehr ihre Gedankengänge prozessbegleitend zu äußern, sondern vielmehr eine nachträgliche Kurzbegründung der vorgenommenen Antwortauswahl zu geben („*Ich hab das angekreuzt, weil . . .*“), möglicherweise weil die ungewohnte Art der Artikulation auf sie befremdlich wirkte, sie sich ihrer Gedankengänge nicht immer bewusst waren oder weil sie diese nicht verbalisieren konnten“ (S. 140, kursiv im Original). Hartmann entschied sich stattdessen für die Durchführung fokussierter Interviews, integrierte dort aber gleich-

wohl Elemente Lauten Denkens, fragte sie doch „*Was geht dir alles durch den Kopf, wenn du diese Frage liest?*“ (S. 141, kursiv im Original).

Ein ähnliches Anliegen verfolgte Richardson in Begleitung der IEA Civic Education Study, mit der „*core concepts of citizenship education*“ erhoben werden sollten (Richardson 2006, S. 164–165). Richardson legte Items des Tests einem Sample von Neuntklässler/-innen vor, um herauszufinden, „*how adolescents interpreted the survey items about political discussion and open classroom climate and the meaning they attached to terms such as ‚discussion‘ and ‚politics*“ (S. 167). Wie Hartmann argumentiert auch Richardson, dass prozessbegleitendes Lautes Denken hierfür nicht geeignet sei, da darin „*specific information about their understanding [...] might not have occurred*“. Sie führte stattdessen ein „*cognitive interview*“ durch, wobei darin „*retrospective questions about thoughts*“ beinhaltet waren (S. 167), sich also wie bei Hartmann Elemente Lauten Denkens wiederfanden.

Die Einschätzungen, ob sich Lautes Denken zur qualitativen Begleitforschung von Tests eignet, variieren somit erheblich und es liegt nahe, für die Diskrepanzen weitere, ausserhalb der Methode liegende Faktoren zu vermuten. Erstens bleibt im Falle Hartmanns (2015), deren Veröffentlichung dem laufenden Forschungsprozess entstammt, letztlich unklar, ob der zugrunde liegende Test die gewünschten Konstrukte überhaupt adressierte. Während Reich den Grund für die Abwesenheit der gewünschten Konzepte in den Items suchte (2009, S. 348), verwarfen allerdings Hartmann und Richardson die Methode. Zweitens ist unklar, inwiefern die Aufgabenformate aufgrund fachlicher Unterschiede eine systematisch unterschiedliche Reichhaltigkeit Lauten Denkens verursacht haben könnten. Auch ist drittens denkbar, dass die Erhebung von Konzepten (Hartmann 2015; Reich 2009; Richardson 2006) schwieriger durch Lautes Denken gelingt, als dies bei Aspekten historischen Denkens zumindest teilweise der Fall war ist (z. B. Ercikan et al. 2015, S. 218). Denkbar ist schliesslich viertens, dass die Güte der für die Analyse eingesetzten theoretischen Konstrukte die differierenden Einschätzungen bewirkten. Während z. B. im Fall von Ercikan et al. (2015) das der Analyse zugrunde liegende Konstrukt *Historical Thinking* differenziert dargelegt wird, finden sich bei Hartmann (2015) allerdings keine Ausführungen zu den eingesetzten Kategorien, die eine Einschätzung erlauben würden. Reich stellt zwar sein Kategoriensystem dar, allerdings ist dieses weitgehend fachunspezifisch, während die interessierenden „*second-order historical concepts*“ gar nicht operationalisiert sind (Reich 2009, S. 351–354). Wird prozessbegleitendes Lautes Denken für die Begleitforschung zu Tests eingesetzt, gilt es also von Fall zu Fall sorgfältig zu prüfen, ob tatsächlich die Methode, der zugrunde liegende Test, die verwendete Theorie oder weitere Faktoren ungeeignet sind, ein erwünschtes Denken oder Verhalten zu evozieren.

Zu (3): Während es bei den bislang genannten Studien um den Umgang mit Materialien oder Tests unter konstruierten Bedingungen ging, wurde Lautes Denken ausserdem vereinzelt eingesetzt, um die Wahrnehmung von ‚natürlichem‘ Unterricht zu erforschen, hier in Form von nachträglichem Lauten Denken. So liess Körber Lernende Aufnahmen von einer zuvor erlebten Unterrichtsstunde betrachten und äussern, was ihnen in den gezeigten Sequenzen „durch den Kopf gegangen sei“ (Körber 2006, S. 199). Auf diese Weise konnte er unter anderem herausfinden, dass

die Schüler/-innen im Falle einer unklaren Adressierung von Operationen zu re-konstruktivem statt de-konstruktivem historischem Denken neigten (S. 211).

Meyer-Hamme interessierte sich dafür, wie „Jugendliche in der deutschen Einwanderungsgesellschaft ihre historischen Identitäten [entwerfen]“ und in welchem Zusammenhang diese mit Geschichtsunterricht stehen (Meyer-Hamme 2009, S. 104) und setzte hierfür biografisch-episodische Interviews sowie das nachträgliche Laute Denken ein. Schüler/-innen sollten anhand von Unterrichtsvideos erzählen, was ihnen „im Laufe der Stunde durch den Kopf gegangen“ sei, konnten hierfür das Video beliebig stoppen und sich dazu äussern. Der Vorteil der Methode liege darin, dass Schüler/-innen „sich konkret und genau beobachtend zu alltäglichen Unterrichtssituationen“ äussern könnten (S. 109).

Zu (4): Ein weiteres Feld, in dem bereits mehrfach Lautes Denken eingesetzt wurde, stellt die Erforschung von Besuchenden an geschichtskulturellen Orten dar, auch hier vor allem mit Blick auf Ausprägungen historischen Denkens. Etwa zielte Baron auf „*interpretive processes historians engage in when ‚reading‘ historic buildings*“, auf „*historical thinking about historic buildings and sites*“, und zwar in der Old North Church in Boston, Massachusetts (Baron 2012, Abstract). Etliche weitere Studien nahmen erwachsene Besuchende in historischen Museen ins Visier. Buttke et al. (2014) interessierten sich für die Auseinandersetzung von Besuchenden mit Objekten im Oberhausmuseum Passau. Thoma (2009) widmete sich dem Lernen in der Ausstellung „Foto und Film“ im Deutschen Museum in München. Wise (2011) interessierten die „*sense-making processes*“ von Besuchenden der Ausstellung des Chelsea Jeans Memorials in New York. Thyroff (2020) nahm Besuchende der Ausstellung „14/18. Die Schweiz und der Grosse Krieg“ und deren Aneignungsweisen in den Blick. Sowohl bei Wise wie auch bei Thyroff konnte durch das Laute Denken nicht nur die kognitive Auseinandersetzung der Besuchenden, sondern auch emotionale Bezugnahmen und die Interaktion mit der Ästhetik und räumlichen Struktur der Ausstellung und deren Elementen prozessbegleitend erhellt werden (Thyroff 2020, S. 411–436, 455–493; Wise 2011, S. 74–89, 123–138).

Die genannten Studien liefern erste prozessbezogene Einblicke in Ausprägungen historischen Denkens an geschichtskulturellen Orten, während das Gros der Forschungsarbeiten in diesem Feld mit pro- oder retrospektiven Methoden arbeitet (im Überblick: Thyroff 2020, S. 143–194). Bisherige prozessbegleitend gewonnene Erkenntnisse beziehen sich in diesem Feld allerdings ausschliesslich auf Erwachsene. Mit Schüler/-innen und mit einer Methode, die als *Stimulated Recall* bezeichnet wird, arbeiteten einzig Lenz und Talsnes. Dabei schien es sich indes eher um eine nachträgliche Reflexion als um nachträgliches Lautes Denken zu handeln. Denn die Schüler/-innen sollten nach Besuch des Holocaust Centers in Oslo über den Grad der Erfüllung ihrer zuvor notierten Erwartungen reflektieren (Lenz und Talsnes 2014, S. 97–100, 113).

Zu (5): Ein weiteres Forschungsfeld stellen vereinzelte Studien zum Planungsverhalten von Geschichtslehrpersonen dar. So findet sich die Methode bei Seidenfuß, der Lehrpersonen unterschiedlicher Expertisegrade bei der Planung einer einstündigen Geschichtslektion laut denken liess, jedoch nicht näher auf Methode und Untersuchungsergebnisse eingeht, lediglich vorschlägt, im Nachgang zum Lauten Denken

ein „Validierungsgespräch“ durchzuführen, um zu geäußerten subjektiven Theorien der Lehrpersonen einen „Konsens zwischen Probanden und Versuchsleiter“ herzustellen (Seidenfuß 2002, S. 65, 70).

Ebenfalls mit einer Kombination aus prozessbegleitendem Lauten Denken und nachgängigen Interviews, ergänzt um Videovignetten, arbeitete Litten (2017, S. 178–193). Lautes Denken gilt ihr als „Methode der Wahl“ für die Erforschung von Planungsverhalten, im konkreten Fall gehe es indes „weniger um die Repräsentation der kognitiven Prozesse als vielmehr um die Erfassung der Inhalte des Denkens und der gedanklichen Organisation dieser Inhalte“ (S. 88, 179–180) – eine Einschätzung, die auch für weitere bereits erwähnte Studien zutrifft (z. B. Thyroff 2020, S. 139–140; Wise 2011). Litten reflektierte zudem mit den Teilnehmenden über Methode und Setting und formuliert durchaus widersprüchliche Befunde: So hätten alle Teilnehmenden „das Planen nicht als von ihrem alltäglichen Vorgehen abweichend empfunden“, gleichzeitig habe sich im Fall von Hauptschullehrpersonen die Erhebungssituation von ihrem Alltag unterschieden, da diese „kooperativ und langfristig planten. Beides konnte in der Laut-Denk-Untersuchung nicht erfasst werden“ (Litten 2017, S. 204–205).

Zu (6): Einen letzten Forschungsbereich stellen Studien zu subjektiven Theorien und Konzepten von (angehenden) Lehrpersonen dar. Larson (1997) erforschte mittels Lautem Denken bei der Betrachtung von Unterrichtsvignetten die Vorstellungen, die Lehrpersonen der Social Studies von Diskussionen haben. Fenn (2015) interessierte sich für Theorien zum Lehren und Lernen im Geschichtsunterricht und setzte hierzu Stimulated Recall ein, wobei die Methode weniger einer Rekonstruktion von Gedanken zu dienen schien als vielmehr dazu, Studierenden „Skripts bewusst zu machen und aufzubrechen“ und zusammen mit Coaches „über alternative Handlungsmuster“ nachzudenken (S. 521).

Schröer (2015) setzt in seiner Studie zu subjektiven Konzepten gelingenden Geschichtsunterrichts bei angehenden Lehrpersonen dem Namen nach Lautes Denken ein, jedoch in einem thematisch begrenzten und vorstrukturierten Setting. Lautes Denken diente dort als Einstieg in ein Interview und wurde durch ein vorgegebenes Mindmap vorstrukturiert, das als „stummer Impuls zum ‚Lauten Denken‘“ fungieren sollte und Teilnehmende nach „einer hinreichenden Denkzeit“ dazu anregen, „im monologischen Selbstgespräch ihre Assoziationen zu Geschichte, zum Geschichtsunterricht und zu ihrem Rollenverständnis“ zu artikulieren (S. 115, 139).

5 Reichweiten und Limitationen

Der Blick auf die beschriebene Forschungslandschaft zeigt, dass die Bezeichnungen „*Lautes Denken* [oder *Stimulated Recall*] inflationär für zum Teil sehr unterschiedliche Formen der Datenerhebung verwendet“ werden (Stark 2010, S. 60). Neben Studien, in welchen Lautes Denken durchaus in klassischer Weise praktiziert wurde, finden sich diverse Spielarten. Mitunter stehen Reflexionen über eine vorausgegangene Situation im Zentrum (Fenn 2015; Lenz und Talsnes 2014), mitunter werden lediglich kurze Elemente Lauten Denkens in andere Erhebungsformen eingeflochten

(Hartmann 2015; Richardson 2006; Schröer 2015). Zu nennen wären ergänzend schliesslich Studien, in denen Proband/-innen zwar nicht aufgefordert wurden, laut zu denken, aber ihr Vorgehen zu erklären (z. B. Gollin und Nitsche 2019, S. 223; Hodel 2013), was als eine verwandte Erhebungsform betrachtet werden kann. Diversifiziert wird das Feld weiter im Hinblick auf die Auswertungsmethoden, die mit Lautem Denken kombiniert werden, etwa qualitative Inhaltsanalyse (z. B. Thyroff 2020) oder Dokumentarische Methode (Lange 2011).

In der Gesamtschau zeigt sich, dass Lautes Denken in der Geschichtsdidaktik bereits einen etablierten Zugang darstellt, während politikdidaktische Fragestellungen bislang nur vereinzelt mit Lautem Denken verfolgt werden. Ein entscheidendes Entwicklungspotenzial besteht also dahingehend, die Methode auch der Politikdidaktik zu erschliessen, beispielsweise zur Erforschung des Planungsverhaltens von Lehrpersonen. Bestimmte Aspekte politischen Denkens und Lernens dürften mit der Methode indes schwieriger zu erheben sein. Insbesondere entziehen sich all diejenigen Situationen, in welchen Kommunikation und Interaktion essenziell sind, wozu Artikulationen von politischer Urteils- und Handlungskompetenz zählen, einer Erforschung durch prozessbegleitendes Lautes Denken (Lyle 2003, S. 862). Weiter gilt es den Wert Lauten Denkens für die Testvalidierung zu überprüfen, wird die Methode doch in der Geschichtsdidaktik im selben Zusammenhang als nützlich beurteilt.

Insgesamt erweist sich die Methode in der Geschichtsdidaktik als ertragreich. Etwa diene sie in mehrerlei Zusammenhängen der Identifizierung von Struktur- und Prozesselementen historischen Denkens, namentlich im Kontext von Experten-Laien-Settings (z. B. Shreiner 2014; Wineburg 1991, 1998; Wise 2011). Allerdings zeigt sich, dass gewonnene Ergebnisse in der Geschichtsdidaktik bis anhin noch wenig vor dem Hintergrund der formulierten Einwände gegen die Methode (Reaktivität und Non-Veridikalität) diskutiert, eingeordnet oder gar relativiert werden. Systematische oder gar systematisch-unterschiedliche Einschränkungen Lauten Denkens in der Auseinandersetzung mit Medientypen wie Text- und Bildquellen oder andere Einschränkungen werden zwar im Methodenteil von Studien benannt (z. B. Lange 2011, S. 113; Litten 2017, S. 179–182), aber nur selten bei der Interpretation der Ergebnisse erneut berücksichtigt (Thyroff 2020, S. 285, 625–626).

Bestätigung findet die in der Methodenliteratur genannte Vielschichtigkeit Lauten Denkens. So treten nicht nur kognitive Aspekte historischen Denkens, sondern auch Wahrnehmungen, Emotionen und Metakognitionen beim Lauten Denken zutage (z. B. Thyroff 2020, S. 455–540, 619–626; Wise 2011). Lautes Denken macht damit dezidiert auch die „nicht-nur-kognitive[n] Momente“ (Borries 2014, S. 167) historischen Denkens und Lernens beobachtbar (Thyroff 2020, S. 619–626). Vereinzelt werden hingegen Schwierigkeiten berichtet, übergreifende Linien historischen Denkens, das zumindest idealtypisch als Abfolge von Teilschritten gedacht wird (z. B. Hasberg und Körber 2003), zu entdecken (Thyroff 2020, S. 617–619). Beim Lauten Denken sollen Teilnehmende auf das reagieren, was sie im unmittelbaren Augenblick sehen, lesen oder erleben (Wineburg 1991, S. 75; Wolfe und Goldman 2005, S. 467) oder tun dies – so beim Ausstellungsbesuch – von sich aus (Thyroff 2020). Das Entwickeln übergreifender Deutungen und Erklärungen wird, im Gegensatz zu Interviews, nicht explizit unterstützt oder durch Instruktionen sogar ausdrücklich

verhindert. Im Umkehrschluss erlaubt Lautes Denken unmittelbare Einblicke in die ganz konkreten Situationen, die Aspekte historischen Denkens evozieren. Die Momentbezogenheit der Methode erweist sich somit als Chance und Begrenzung gleichermaßen (Thyroff 2020, S. 617–619). Ertragreich scheinen vor diesem Hintergrund insbesondere Studien, die Lautes Denken um weitere Erhebungsformen wie vertiefende Interviews ergänzen.

Lautes Denken eignet sich, so lässt sich zusammenfassen, für all diejenigen Forschungsanliegen, in welchen unmittelbare Einsichten in kognitive Prozesse von Interesse sind, steht innerhalb dessen aber einer grossen Bandbreite von Forschungsfragen und -designs offen. Angesichts vielfältiger Einsatzmöglichkeiten Lauten Denkens erscheint es umso dringlicher, das jeweilige im konkreten Fall gewählte Forschungsdesign differenziert darzulegen und zu begründen und gewonnene Ergebnisse vor dem Hintergrund der Methode und ihrer Einschränkungen zu reflektieren.

Literatur

- Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedien. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Strategien und Regulationsaktivitäten*. Münster: Waxmann.
- Baron, C. (2012). Understanding historical thinking at historic sights. *Journal of Educational Psychology*, 104(3), 833–847. <https://doi.org/10.1037/a0027476>.
- Bernstein, J. F. (2010). Using <Think alouds> to Understand Variations in Political Thinking. *Journal of Political Science Education*, 6(1), 49–69. <https://doi.org/10.1080/15512160903467695>.
- Borries, B. von. (2014). *Zwischen „Genuss“ und „Ekel“: Ästhetik und Emotionalität als konstitutive Momente historischen Lernens (mit Beiträgen von Johannes Meyer-Hamme)*. Schwalbach/Ts.: Wochenschau.
- Buttkereit, F.-A., Heinrich, H.-A., Reithmayr, L., Watmann, A., & Weyland, V. (2014). *Prozeßbegleitendes lautes Denken im Museum: Methodenbericht und Ergebnisdokumentation [Internet-Dokument]*. <https://opus4.kobv.de/opus4-uni-passau/frontdoor/index/index/docId/229>. Zugegriffen am 04.05.2021.
- Cho, B.-Y, Han, H., & Kucan, Linda L. (2018). An Exploratory Study of Middle-School Learners' Historical Reading in an Internet Environment. *Reading and Writing*, 31(7), 1525–1549. <https://doi.org/10.1007/s11145-018-9847-4>.
- Deffner, G. (1984). *Lautes Denken. Untersuchung zur Qualität eines Datenerhebungsverfahrens*. Bern: Lang.
- Ercikan, K., Seixas, P., Lyons-Thomas, J., & Gibson, L. (2015). Cognitive validity evidence for validating assessments of historical thinking. In K. Ercikan & P. Seixas (Hrsg.), *New directions in assessing historical thinking* (S. 206–220). New York: Routledge.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis. Verbal reports as data*. Cambridge: MIT.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data*. Cambridge: MIT.
- Fenn, M. (2015). Beeinflusst geschichtsdidaktische Lehre die subjektiven Theorien von Studierenden zu Lehren und Lernen im Geschichtsunterricht? Ergebnisse einer empirischen Interventionsstudie. *Geschichte in Wissenschaft und Unterricht*, 66(9/10), 515–538.

- Fromann, U. (2005). *Die Methode „Lautes Denken“*. https://www.e-teaching.org/didaktik/qualitaet/usability/Lautes%20Denken_e-teaching_org.pdf. Zugegriffen am 04.05.2021.
- Gollin, K., & Nitsche, M. (2019). Schreibprozesse in Geschichte. In M. Waldis & B. Ziegler (Hrsg.), *Forschungswerkstatt Geschichtsdidaktik 17. Beiträge zur Tagung „geschichtsdidaktik empirisch 17“* (S. 218–231). Bern: hep.
- Hartmann, M. (2015). Lässt sich konzeptuelles Wissen eigentlich messen? In A. Petrik (Hrsg.), *Formate fachdidaktischer Forschung in der politischen Bildung* (S. 136–145). Schwalbach/Ts.: Wochenschau.
- 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.
- Hodel, J. (2013). *Verkürzen und Verknüpfen. Geschichte als Netz narrativer Fragmente: Wie Jugendliche digitale Netzmedien für die Erstellung von Referaten im Geschichtsunterricht verwenden*. Bern: hep.
- Houwen, A., Boxtel, C. v., & Holthuis, P. (2020). Dutch Students' Understanding of the Interpretative Nature of Textbooks When Comparing Two Texts about a Significant Event in the Development of Democracy. *History Education Research Journal*, 17(2), 214–28. <https://doi.org/10.14324/HERJ>.
- Iordanou, K., Kendeou, P., & Zembylas, M. (2020). Examining My-Side Bias during and after Reading Controversial Historical Accounts. *Metacognition and Learning*, 15(3), 319–342. <https://doi.org/10.1007/s11409-020-09240-w>.
- Jay, L. (2021). Revisiting Lexington Green: Implications for Teaching Historical Thinking. *Cognition and Instruction*, Februar, 1–22. <https://doi.org/10.1080/07370008.2021.1880410>.
- Kaiser, R. (2018). Das Konzept Metakognition. In A. Kaiser, R. Kaiser, A. Lambert, K. Hohenstein & R. Kaiser (Hrsg.), *Metakognition: Die Neue Didaktik* (S. 31–68). Göttingen: V&R. <https://doi.org/10.13109/9783666702549.31>.
- Kaliski, P., Smith, K., & Huff, K. (2015). The Importance of Construct Validity Evidence in History Assessment: What Is Often Overlooked or Misunderstood? In K. Ercikan & P. Seixas (Hrsg.), *New directions in assessing historical thinking* (S. 195–205). New York: Routledge.
- Konrad, K. (2010). Lautes Denken. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 476–490). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Körber, A. (2006). Hinreichende Transparenz der Operationen und Modi des historischen Denkens im Unterricht? Erste Befunde einer Einzelfallanalyse im Projekt „FUER Geschichtsbewusstsein“. In H. Günther-Arndt & M. Sauer (Hrsg.), *Geschichtsdidaktik empirisch. Untersuchungen zum historischen Denken und Lernen* (S. 189–214). Münster: LIT.
- Lange, K. (2011). *Historisches Bildverstehen oder Wie lernen Schüler mit Bildquellen? Ein Beitrag zur geschichtsdidaktischen Lehr-Lern-Forschung*. Münster: LIT.
- Larson, B. E. (1997). Social Studies Teacher's Conceptions of Discussion: A Grounded Theory Study. *Theory & Research in Social Education*, 25(2), 113–136. <https://doi.org/10.1080/00933104.1997.10505798>.
- Lenz, C., & Talsnes, A. (2014). Stimulated recall. Tracing learning narratives and reinforcing historical competence. In H. Bjerg, A. Körber, C. Lenz & O. Wrochem (Hrsg.), *Teaching historical memories in an intercultural perspective: Concepts and methods: Experiences and results from the TeacMem project* (S. 97–114). Berlin: Metropol.
- Litten, K. (2017). *Wie planen Geschichtslehrkräfte ihren Unterricht? Eine empirische Untersuchung der Unterrichtsvorbereitung von Geschichtslehrpersonen an Gymnasien und Hauptschulen*. V&R unipress. <http://www.vr-elibrary.de/doi/book/10.14220/9783737006972>. Zugegriffen am 04.05.2021.
- Logtenberg, A. (2012). *Questioning the past: Student questioning and historical reasoning*. Amsterdam: Ipskamp.
- Lyle, J. (2003a). Stimulated recall: a report on its use in naturalistic research. *British Educational Research Journal*, 29(6), 861–878.

- Maag Merki, K., & Werner, S. (2014). Erfassung und Bewertung professioneller Kompetenz von Lehrpersonen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 745–763). Münster: Waxmann.
- McGrew, S. (2021). Internet or Archive? Expertise in Searching for Digital Sources on a Contentious Historical Question. *Cognition and Instruction*, 4, 1–29. <https://doi.org/10.1080/07370008.2021.1908288>.
- Meyer-Hamme, J. (2009). *Historische Identitäten und Geschichtsunterricht. Fallstudien zum Verhältnis von kultureller Zugehörigkeit, schulischen Anforderungen und individueller Verarbeitung*. Idstein: Schulz-Kirchner.
- Reich, G. A. (2009). Testing historical knowledge: standards, multiple-choice questions and student reasoning. *Theory & Research in Social Education*, 37(3), 325–360. <https://doi.org/10.1080/00933104.2009.10473401>.
- Richardson, W. K. (2006). Combining cognitive interviews and social science surveys: strengthening interpretation and design. In K. C. Barton (Hrsg.), *Research methods in social studies education* (S. 159–182). Charlotte: Information Age.
- Riegel, U. (2013). Videobasierte Kompetenzforschung in den Fachdidaktiken. Einleitung. In U. Riegel & K. Macha (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (S. 9–24). Münster: Waxmann.
- Rogiers, A., Merchie, E., & Van Keer, H. (2020). What they say is what they do? Comparing task-specific self-reports, think-aloud protocols, and study traces for measuring secondary school students' text-learning strategies. *European Journal of Psychology of Education*, 35, 315–332.
- Saaman, D. (2001). Kognitive Prozeßvariablen und Lautes Denken: Individuelle Unterschiede in der Verbalisierung. In J. F. Schneider (Hrsg.), *Lautes Denken. Prozessanalysen bei Selbst- und Fremdeinschätzungen* (S. 125–141). Weimar: Rita Dadder.
- Schnell, C. (2016). „Lautes Denken“ als qualitative Methode zur Untersuchung der Validität von Testitems. Erkenntnisse einer Studie zur Diagnose des ökonomischen Fachwissens von Schülerinnen und Schülern der Sekundarstufe I. *Zeitschrift für ökonomische Bildung*, 5, 26–49.
- Schröer, L. (2015). *Individuelle didaktische Theorien und Professionswissen: Subjektive Konzepte gelingenden Geschichtsunterrichts während der schulpraktischen Ausbildung*. Münster: LIT.
- Seidenfuß, M. (2002). Was denken erfolgreiche Geschichtslehrer bei der Unterrichtsplanung? Expertiseansatz und empirische Forschung in der Didaktik der Geschichte. In S. Handro & B. Schönemann (Hrsg.), *Methoden geschichtsdidaktischer Forschung* (S. 61–71). Münster: LIT.
- Shreiner, T. L. (2014). Using historical knowledge to reason about contemporary political issues: An expert – novice study. *Cognition and Instruction*, 32(4), 313–352. <https://doi.org/10.1080/07370008.2014.948680>.
- Shreiner, T. L. (2019). Students' use of data visualizations in historical reasoning: A think-aloud investigation with elementary, middle, and high school students. *The Journal of Social Studies Research*, 43(4), 389–404. <https://doi.org/10.1016/j.jssr.2018.11.001>.
- Smith, M. D., Breakstone, J., & Wineburg, Sam (2019). History Assessments of Thinking: A Validity Study. *Cognition and Instruction*, 37(1), 118–144. <https://doi.org/10.1080/07370008.2018.1499646>.
- Someren, M. W. van, Barnard, Y. F., & Sandberg, Jacobijn A. C. (1994). *The think aloud method. A practical guide to modelling cognitive processes*. Cambridge: Academic Press.
- Stark, T. (2010). Lautes Denken in der Leseprozessforschung. Kritischer Bericht über eine Erhebungsmethode. *Didaktik Deutsch*, 15(29), 58–63.
- Thoma, G.-B. (2009). *Was lernen Besucherinnen und Besucher im Museum? Eine Untersuchung von Lerngelegenheiten einer Museumsausstellung und ihrer Nutzung*. http://macau.uni-kiel.de/servlets/MCRFileNodeServlet/dissertation_derivate_00002960/dissertation_gun-brit_thoma.pdf. Zugriffen am 04.05.2021. [besucht: 17.12.2020]
- Thyroff, J. (2020). *Aneignen in einer historischen Ausstellung. Eine Bestandsaufnahme von Elementen historischen Denkens bei Besuchenden der Ausstellung „14/18. Die Schweiz und der Grosse Krieg“*. Bern: hep. <https://doi.org/10.36933/9783035516258>.

- Wallach, D., & Wolf, C. (2001). Das prozeßbegleitende Laute Denken. Grundlagen und Perspektiven. In J. F. Schneider (Hrsg.), *Lautes Denken. Prozessanalysen bei Selbst- und Fremdeinschätzungen* (S. 9–29). Weimar: Rita Dadder.
- Weidle, R., & Wagner, A. C. (1982). Die Methode des Lauten Denkens. In G. L. Huber & H. Mandl (Hrsg.), *Verbale Daten: Eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung* (S. 81–103). Weinheim: Beltz.
- Wineburg, S. 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. S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science*, 22(3), 319–346. https://doi.org/10.1207/s15516709cog2203_3.
- Wineburg, S. S. (2001). *Historical thinking and other unnatural acts. Charting the future of teaching the past*. Philadelphia: Temple UP.
- Wise, S. (2011). *Visitors encounter the dust: How People think with objects in a history museum exhibition*. [InternetDocument]. <https://purl.stanford.edu/zs034ry4864>. Zugegriffen am 04.05.2021. [besucht: 17.12.2020]
- Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, 23(4), 467–502.

Kommentierte Literatur

Vielzitierte „Klassiker“ der Methodenliteratur zum Lauten Denken sind die Publikationen von Ericsson und Simon aus den Jahren 1980, 1984 und 1993, unter anderem:

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data*. Cambridge: MIT.

Als aktuellere und knappere, aber gleichwohl differenzierte Einführung in die Methode, ihre Potenziale und Herausforderungen eignen sich die Texte von Konrad und Stark, die beide über ihre Bezugsdisziplinen hinaus lesenswert sind:

Konrad, K. (2010b). Lautes Denken. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 476–490). Wiesbaden: VS Verlag für Sozialwissenschaften.

Stark, T. (2010b). Lautes Denken in der Leseprozessforschung. Kritischer Bericht über eine Erhebungsmethode. *Didaktik Deutsch*, 15(29), 58–63.

Als spezifische Einführung in die Spielart *Stimulated Recall* kann zudem Lyle dienen:

Lyle, J. (2003b). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29(6), 861–878.