

Historische Kompetenzen mit offenen Antwortformaten messen – Eine Studie auf Basis der »Sechser-Matrix« des FUER-Modells

Christiane Bertram, Wolfgang Wagner, Elisabeth Schaser

1. Einleitung

Seit der »kopernikanischen Wende« (Frederking 2008, 5) in der **Bildungssteuerung** – gekennzeichnet durch den Paradigmenwechsel von der Input- **zur Output-Orientierung** (Klieme et al. 2003) und die Bedeutung des Kompetenzerwerbs (Weinert 2001) – werden auch in Fächern wie Geschichte, Politik oder Ethik, die »schwer messbare Kompetenzen« (Frederking 2008) adressiert und empirisch abgesicherte Kompetenzmodelle und **Messinstrumente** entwickelt, so z. B. im DFG-Projekt »Literarästhetische Urteilskompetenz (LUK)« (Frederking/Roick/Steinhauer 2010) oder im vom BMBF geförderten Projekt »HiTCH = Historical Thinking – Competencies in History«. Der vorliegende Beitrag zur **Entwicklung und Validierung eines Messinstruments mit offenen Aufgabenformaten auf Basis des historischen Kompetenz-Strukturmodells** der FUER-Gruppe¹ (Körber/Schreiber/Schöner 2007), **genauer der sogenannten »Sechser-Matrix«** (vgl. Abbildung 1), ist in diesem Kontext zu verorten.

¹ Die Gruppe »FUER Geschichtsbewusstsein« hat sich zum Ziel gesetzt, die Förderung eines reflektierten und (selbst-) reflexiven Geschichtsbewusstseins durch Grundlagenforschung und Empirie zu präzisieren.

2. Theoretischer Hintergrund

2.1 Historisches Kompetenz-Strukturmodell der FUER-Gruppe

Das FUER-Modell fasst, im Anschluss an Rüsen (1983), historisches Denken als einen in einem Regelkreis modellierbaren Prozess auf: Verunsicherungen und Interessen setzen den Prozess historischen Denkens in Gang, der sich – in einer Fragestellung gebündelt – entweder in re-konstruierender Absicht an die Vergangenheit richtet oder sich in de-konstruierender Absicht mit vorliegenden historischen Narrationen auseinandersetzt. Das Ergebnis, das sich als eigene Narration bzw. als Stellungnahme gegenüber einer Darstellung präsentiert, befriedigt entweder bereits die historischen Orientierungsbedürfnisse oder führt zu einer neuen historischen Frage. Durch den an verschiedenen Themen und Fragestellungen immer wieder durchlaufenen Prozess historischen Denkens bilden sich historische Sachkompetenzen heraus, das heißt, die Schülerinnen und Schüler verfügen in zunehmenden Maße über die für den Umgang mit Geschichte relevanten Prinzipien, Konzepte und Skripts (Schreiber et al. 2007).

2.2 Historische Methodenkompetenzen und Fokussierungen

Ausgehend von einem narrativistischen Geschichtsverständnis (Danto 1980; Ricœur 1988; Rüsen 1983) werden im FUER-Modell unter den »historischen Methodenkompetenzen« Re- und De-Konstruktionskompetenzen gefasst, die sich jeweils auf die Fokussierungen »Vergangenes«, »Geschichte« und »Gegenwart/Zukunft« beziehen. In der Re-Konstruktion (im Folgenden: »Re-Ko«) werden einer Fragestellung folgend mithilfe der Quellenkritik und -interpretation Vergangenheitspartikel erhoben (Fokussierung auf Vergangenes). Diese Partikel werden unter Berücksichtigung des synchronen und diachronen Kontextes dem Medium und Adressaten entsprechend erzählt (Fokussierung auf Geschichte), wodurch – bewusst oder unbewusst, subtil oder direkt – eine Orientierung für heute und morgen gegeben wird (Fokussierung auf Gegenwart und Zukunft). Die De-Konstruktion (im Folgenden »De-Ko«) analysiert – ebenfalls bezogen auf die drei Fokussierungen – die Tiefenstruktur einer vorliegenden historischen Narration (Schreiber et al. 2007). Die Operationen der Re- und De-Konstruktionskompetenzen werden auf die drei Fokussierungen bezogen, sodass eine Sechs-Felder-Matrix entsteht (vgl. Abbildung 1).

2.3 Empirische und narrative Triftigkeit

Das Grundmuster einer historischen Narration besteht darin, dass mindestens zwei verschiedene, zeitlich differente Ereignisse sinnhaft miteinander verknüpft werden, sodass eine meist sprachlich vermittelte Verlaufsstruktur entsteht (u. a. Barricelli/Gautschi/Körber 2012; Jeismann 1977; Körber Schreiber/Schöner 2007; Pandel 2010; Rüsen 1983). Da die Vergangenheit unwiederbringlich vorbei ist, sind wir immer auf eine (Re-)Konstruktion des Vergangenen angewiesen, deren Plausibilität überprüft werden muss. Als Ankerpunkt für die Analyse einer Narration werden im FUER-Modell – wie auch in anderen historischen Kompetenzmodellen (Barricelli/Gautschi/Körber 2012) – die von Rüsen (1983; 2013) entwickelten »Triftigkeiten« genutzt (Schreiber et al. 2007; von Borries 2009). Die »Wahrheitskriterien des historischen Denkens« (Rüsen 1983, 76) zeigen sich u. a. auf der empirischen und der narrativen Ebene. Empirisch triftig sind Geschichten, wenn die »behaupteten Tatsachen durch Erfahrungen gesichert sind« (Rüsen 1983, 82). Daher umfasst die Überprüfung der empirischen Triftigkeit einer Narration die Untersuchung der Vergangenheitspartikel wie auch der Kontextualisierung bezogen u. a. auf ihren Wahrheitsgehalt (Kühberger 2009). Die Ebene der »narrativen Triftigkeit« hingegen zeigt sich in der »Einheit einer in sich sinnvollen Erzählung«, deren Erzählfluss durch ein »Sinnkriterium« gebildet wird (Rüsen 1983, 83), welches sich in einer logischen und in sich schlüssigen argumentativen Verknüpfung der Inhalte ausdrückt (Kühberger 2009).

2.4 Fragestellung der Studie

Im folgenden Beitrag wird die Fragestellung verfolgt, ob sich die Sechser-Matrix dazu eignet, offene Aufgabenformate zu entwickeln wie auch theorieentsprechend zu codieren und zu skalieren. Wenn die in den offenen Formaten gemessenen Schülerleistungen mit weiteren Leistungsmaßen, die mit geschlossenen Skalen erfasst wurden, korrelieren, wäre dies ein Hinweis darauf, dass mit den offenen Aufgaben historische Kompetenzen erfasst werden.

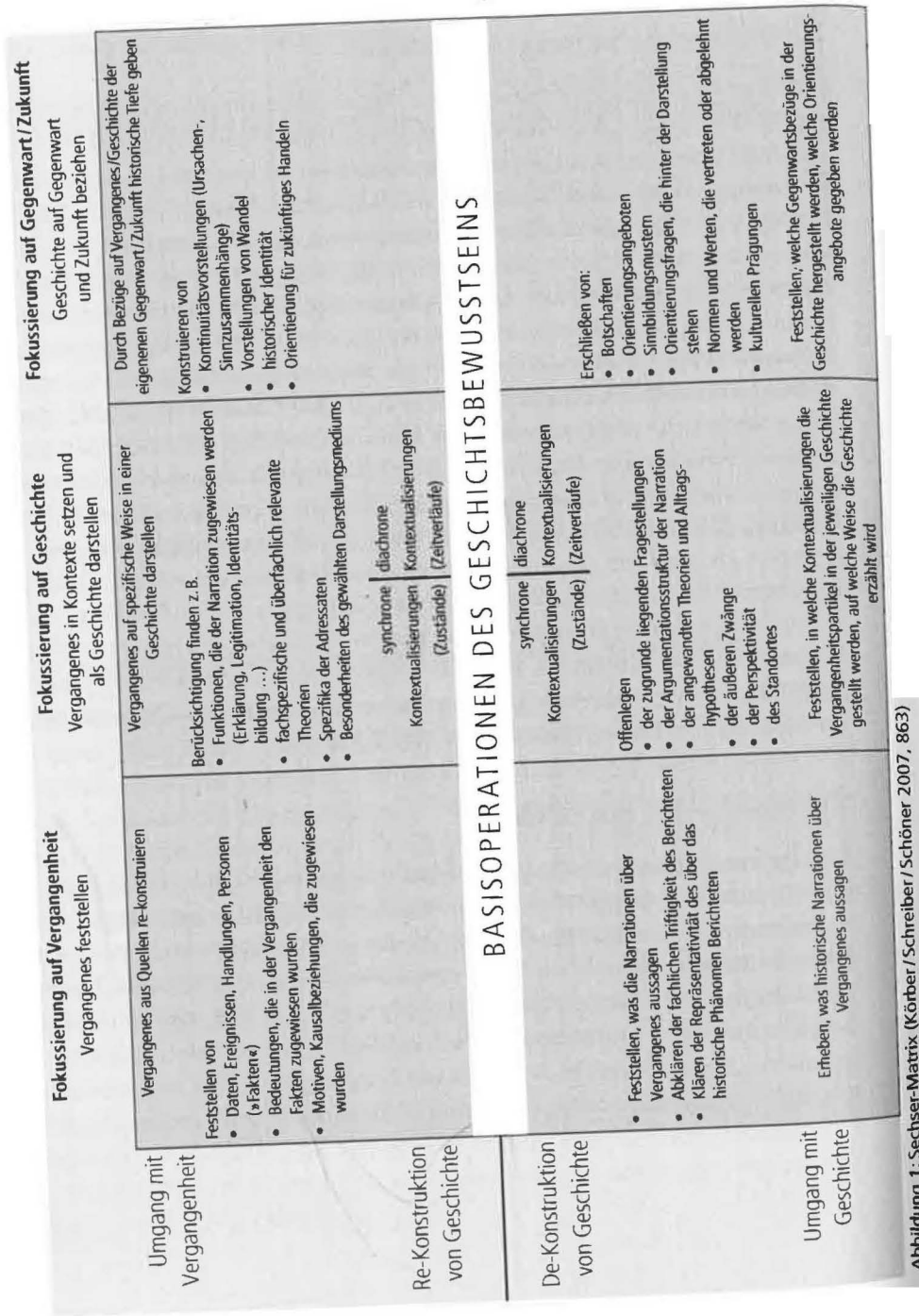


Abbildung 1: Sechser-Matrix (Körber/Schreiber/Schröder 2007, 863)

3. Methode

3.1 Stichprobe

Die Daten wurden anlässlich einer vom BMBF geförderten Interventionsstudie zur Wirksamkeit von Zeitzeugenbefragungen erhoben. Nach Abschluss der Intervention bearbeiteten die Schülerinnen und Schüler u. a. **sieben offene Aufgaben**, in denen sie Materialien zur Demonstration am 9. Oktober 1989 de- und rekonstruierten. Der gesamte Datensatz besteht aus Schülerantworten von 38 neunten Gymnasialklassen ($N = 962$). Bisher wurden zu den sieben offenen Aufgaben 190 Schülerantworten ausgewertet.

3.2 Messinstrumente im Nachttest

Im Nachttest wurden offen und geschlossen formulierte Kompetenzaufgaben eingesetzt. Die offenen Aufgaben bezogen sich auf drei Texte zum 9. Oktober 1989 (Schulbuchtext, Aufruf von Kurt Masur u. a. vor der Demonstration und ein Zeitzeugenbericht). Ob die Schülerinnen und Schüler anhand dieser drei Texte zwischen Quellen und Darstellungen unterscheiden können, wurde ebenso wie ihr grundsätzliches Verständnis von Geschichte als Rekonstruktion der Vergangenheit mit geschlossenen Items erhoben. Das Begriffsverständnis wird im FUER-Modell, ebenso wie die Verfügung über die epistemologischen Prinzipien des historischen Denkens, den Sachkompetenzen zugeordnet (Schöner 2007). Im Folgenden wird die Reliabilität der geschlossenen Aufgaben bezogen auf die Gesamtstichprobe ($N = 962$) berichtet.

Tests zu den epistemologischen Überzeugungen. In mehreren Erhebungen wurden die 40 Items faktorenanalytisch überprüft.² Hierbei ließen sich die Daten mit drei Faktoren am besten erklären, sodass sich drei Skalen ergaben. Die Skala »Sachkompetenz ›Re-Ko«« (24 Items, Weighted Likelihood Estimate (WLE) Person Separation Reliability (PSR) = .87³) erfasst, ob die Probandinnen und Probanden über die Einsicht verfügen, dass Geschichte aus dem Vergleich und der Interpre-

2 Zur Entwicklung und Erprobung der Messinstrumente im Überblick vgl. Bertram/Wagner/Trautwein (2013), zur Erprobung des Kurzinstrumentes zur Erfassung der epistemologischen Überzeugungen vgl. Bertram/Wagner/Trautwein (2014).
3 Die Reliabilität eines Kompetenztests kann nach der Item-Response-Theorie mit der Maßzahl Weighted Likelihood Estimate (WLE) Person Separation Reliability (PSR) angegeben werden, die ähnlich wie Cronbachs Alpha zu interpretieren ist (Faustregel: > .70 gute Reliabilität).

tation von Quellen und Darstellungen entsteht. Ob es den Schülerinnen und Schülern bewusst ist, dass historische Narrationen de-konstruiert werden müssen, wird mit der Skala »Sachkompetenz ›De-Ko« (6 Items, WLE PSR = .59⁴) gemessen. Die dritte Skala »Sachkompetenz ›Zeitzeuge« (10 Items, WLE PSR = .73) erfasst, ob die Schülerinnen und Schüler die Besonderheit von Zeitzeugenberichten, z. B. im Hinblick auf ihre (Retro-) Perspektivität und Selektivität, verstanden haben.

Test zur Verfügung über die Konzepte »Quelle« und »Darstellung«. Zu den drei Texten (Schulbuchtext, Aufruf, Zeitzeugenbericht) wurden Aussagen hinsichtlich ihrer Einordnung als Quelle oder Darstellung formuliert, die die Schülerinnen und Schüler auf ihren Wahrheitsgehalt hin einordnen sollten (23 Items, WLE PSR = .74)

3.3 Konstruktion der offenen Aufgabenformate

Die sechs Felder der Sechser-Matrix wurden in jeweils eine Aufgabe, im Feld »Re-Ko Vergangenes« in zwei Aufgaben, übersetzt.⁵ Die ersten drei Aufgaben beziehen sich auf einen Schulbuchtext zum 9. Oktober 1989, den die Schülerinnen und Schüler im Hinblick auf die drei Fokussierungen de-konstruierten (vgl. Tabelle 1).

Tabelle 1: Übersicht der De-Konstruktionsaufgaben

Aufgabe	Material	Kompetenz-adressierung	Aufgabenformulierung (gekürzt)
1	Schulbuch	De-Konstruktion, Fokus »Vergangenes«	Welche Fakten zur Montagsdemonstration hat der Autor ausgewählt?
2	Schulbuch	De-Konstruktion, Fokus »Geschichte«	Liste die Aussagen auf, in denen der Autor Interpretationen vornimmt und erkläre, wie er in diesen Aussagen die Ereignisse deutet!
3	Schulbuch	De-Konstruktion, Fokus »Gegenwart/Zukunft«	Welche Schlüsse sollen die Schülerinnen und Schüler für ihre Zukunft aus dem Schulbuchtext ziehen?

4 Die niedrige Reliabilität der Skala lässt sich durch die Schwierigkeit der Items für Schülerinnen und Schüler der 9. Jahrgangsstufe erklären. Bei Studierenden weist die De-Ko-Skala deutlich höhere Reliabilitätswerte auf (Bertram / Wagner / Trautwein 2014).

5 Prof. Dr. Waltraud Schreiber und Stefanie Zabold unterstützten uns bei der Aufgabenformulierung und der Entwicklung des Codesystems.

Danach bekamen die Schülerinnen und Schüler zwei weitere Quellen. Aus dem »Aufruf an alle Leipziger«, der von Kurt Masur u. a. direkt vor der Demonstration verlesen wurde, wie auch aus dem Zeitzeugenbericht von Christian Führer, damals Pfarrer an der Nikolaikirche, und aus dem Schulbuchtext sollten sie die Ereignisse des 9. Oktober rekonstruieren (vgl. Tabelle 2).

Tabelle 2: Übersicht über die De-Konstruktionsaufgaben

Aufgabe	Material	Kompetenz-adressierung	Aufgabenformulierung (gekürzt)
4	Aufruf	Re-Konstruktion, Fokus »Vergangenes«	Welche weiteren Informationen zur Situation vor und zum Verlauf der Demonstration werden in den beiden Texten gegeben?
5	Zeitzeugenbericht	Re-Konstruktion, Fokus »Vergangenes«	
6	Alle Materialien	Re-Konstruktion, Fokus »Geschichte«	Wie erklärst du dir, wenn du alle Informationen zusammennimmst, dass die Demonstration am 9. Oktober friedlich verlief?
7	Alle Materialien	Re-Konstruktion, Fokus »Gegenwart/Zukunft«	In der Schulbuchredaktion entsteht ein Streit, ob in der Neuauflage die Darstellung der Demonstration vom 9. Oktober 1989 aus Platzgründen ganz gestrichen werden soll. Sammle Argumente und begründe, warum weiterhin über die Demonstration berichtet werden soll.

3.4 Codierung und Skalierung

Die Codierung der offenen Schülerantworten wurde in MAXQDA 10 vorgenommen und orientierte sich an der Methode der qualitativen Inhaltsanalyse nach Mayring (2010), wobei die vorliegenden Schülertexte, »the manifest content of the communication« (Berelson 1952, 16), direkt codiert wurden. Der Codebaum wurde deduktiv aus der Aufgabenstellung, wie auch induktiv, »in vivo« (Kuckartz 2007, 74), aus den Schülerantworten entwickelt und in Memos erläutert. Die Codierungen wurden als Variablen nach IBM SPSS Statistics 20 exportiert und skaliert.

Für jede der sieben Aufgaben wurde ein differenzierter Codebaum entwickelt⁶, der im Rahmen dieses Beitrags nicht im Detail dargestellt werden kann. In der Aufgabe 1 wurde pro Fakt ein Punkt vergeben, wenn mindestens zwei Bestandteile eines Vergangenheitspartikels, z. B. Agens und Aktion, genannt wurden, um unverbundene Informationsetzen (z. B. »90 000 Demonstranten«) aus der Wertung auszuschließen. In der Aufgabe 2 wurde für jedes korrekt genannte Zitat und jede zutreffende Deutung jeweils ein Punkt vergeben. In der Aufgabe 3 wurden textbasierte Botschaften (z. B. »Heute gilt der Tag als letzter Tag der DDR, weil an diesem Tag das Demonstrationsrecht durchgesetzt wurde.«) wie auch triftige individuelle Schlussfolgerungen (z. B. »Kampf für Menschenrechte«) mit jeweils einem Punkt gewertet. In den Aufgaben 4 und 5 wurde für jede korrekt genannte Information ein Punkt vergeben. Bei der Bewertung der Aufgabe 6 wurden a) die Summe der genannten Erklärungen, b) die Anzahl der Informationsquellen und c) das argumentative Niveau (niedrig – mittel – hoch) berücksichtigt. In der Aufgabe 7 wurden a) die Anzahl der Argumente und b) das argumentative Niveau in die Wertung einbezogen.

Am Beispiel von Aufgabe 6 sollen das Vorgehen bei der Codierung und die theoriegeleitete Skalierung kurz vorgestellt werden. Da in dieser wie auch in Aufgabe 7 die Schülerinnen und Schüler re-konstruierten, wurden die Texte hinsichtlich der empirischen und narrativen Triftigkeit codiert. Die Re-Konstruktionsleistung der Schülerinnen und Schüler zeigte sich darüber hinaus an der Verwendung mehrerer Informationsquellen.

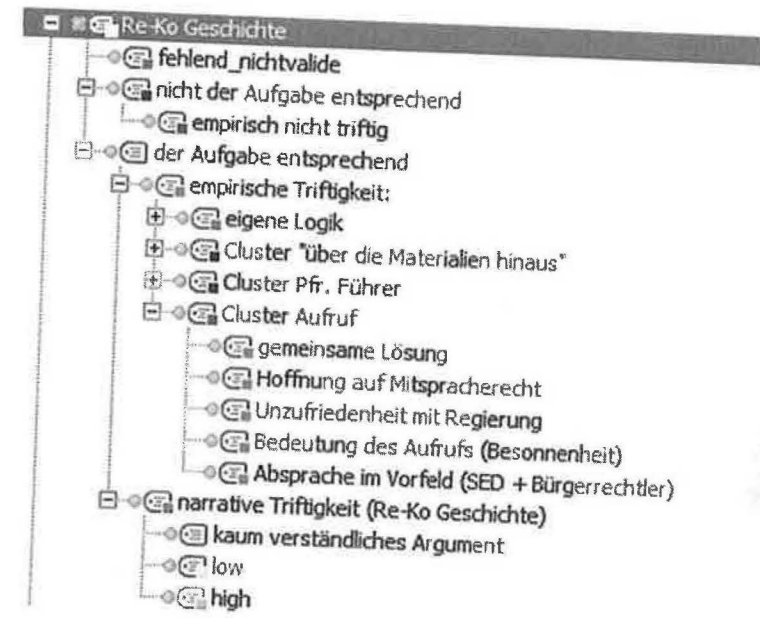


Abbildung 2: Auszug aus dem Codesystem für Aufgabe 6

Wie der Auszug aus dem Codesystem zeigt (vgl. Abbildung 2), wurden empirisch triftige Angaben im jeweiligen Cluster mit dem passenden Subcode gekennzeichnet. Zudem wurde die Antwort hinsichtlich des argumentativen Niveaus codiert. Diese drei Ebenen (Anzahl der empirisch triftigen Angaben, Anzahl der genutzten Cluster und das sprachliche Niveau) wurden durch die Berechnung des geometrischen Mittels, das qualitativ unterschiedliche Dimensionen angemessen berücksichtigen kann, in einer Maßzahl zusammengeführt.⁷

6 Im Rahmen ihrer Bachelor-Arbeit hat Elisabeth Schaser in enger Absprache mit der Erstautorin den Codebaum entwickelt und das Teilsample codiert. 20% des Samples ($N = 34$) wurden später von der Erstautorin codiert. Die in MAXQDA berechnete Inter-coder-Übereinstimmungen bezogen auf die Schülerdokumente liegen zwischen 50% und 90% (MW: 79,5%, SD: 9,54).

7 Die in den zwei bzw. drei Dimensionen vergebenen Punkte können nicht einfach aufsummiert und gemittelt werden, da das Verhältnis zwischen den zwei bzw. drei Dimensionen betrachtet werden muss. Daher eignet sich hier das geometrische Mittel, bei dem die n -te Wurzel aus dem Produkt der Messwerte gezogen wird, als Maßzahl (vgl. <http://www.mathe-lexikon.at/statistik/lagemasse/mittelwert/geometrisches-mittel.html> [12.09.2014]).

4. Ergebnisse

Auf Basis unterschiedlicher quantitativer Auswertungen (Korrelations- und Faktorenanalysen) wurden Zusammenhänge zwischen den in SPSS berechneten Leistungsscores in den sieben offenen Aufgaben mit weiteren Leistungsmaßen untersucht.

4.1 Deskriptive Ergebnisse

In Tabelle 3 werden bezogen auf die Teilstichprobe ($N = 190$) die zentralen deskriptiven Daten der in den folgenden Analysen verwendeten Tests zusammengestellt.

Tabelle 3: Übersicht über die deskriptiven Daten der Testinstrumente

Geschlossene Testformate	Beschreibung	M	SD
Sachkompetenz »Re-Ko«	Einsicht, dass Geschichte aus der Vergangenheit re-konstruiert wird	0.09	1.06
Sachkompetenz »De-Ko«	Einsicht, dass Narrationen de-konstruiert werden müssen	0.04	1.17
Sachkompetenz »Eigenart des Zeitzeugen«	Einsicht in die Eigenart des Zeitzeugen	0.09	0.95
Sachkompetenz »Quelle oder Darstellung?«	Begriffskennntnis hinsichtlich der Unterscheidung von Quellen und Darstellungen	0.06	0.68
Offene Testformate	Beschreibung	M	SD
A1: De-Ko Vergangenes	Fakten im Schulbuchtext	5.05	3.47
A2: De-Ko Geschichte	Deutungen im Schulbuchtext	2.49	1.89
A3: De-Ko Gegenwart/Zukunft	Botschaft des Schulbuchtexts	1.51	0.98
A4: Re-Ko Vergangenes (Aufruf)	Informationen im Aufruf	2.20	1.59
A5: Re-Ko Vergangenes (Zeitzeuge)	Informationen im Zeitzeugenbericht	3.28	1.85
A6: Re-Ko Vergangenes Geschichte	Erklärung des friedlichen Verlaufs	2.11	0.83
A7: Re-Ko Vergangenes Gegenwart/Zukunft	Warum weiter über 9.10.1989 berichten?	1.71	1.15

4.2 Interkorrelationen der offenen Aufgabenformate

Nachdem auf der Basis von Häufigkeitsanalysen sichergestellt war, dass es weitgehend keine Boden- oder Deckeneffekte gab⁸, wurden im nächsten Schritt Korrelationen zwischen den offenen Aufgaben berechnet. Aufgabe 3 korrelierte mit keiner anderen Aufgabe statistisch signifikant. Aufgabe 1 zeigte mit allen Aufgaben (außer mit Aufgabe 3) statistisch signifikante Zusammenhänge ($.19 \leq r \leq .35$). Die Aufgabe 6 und 7 korrelierten (außer mit den Aufgaben 3 und 4) statistisch signifikant mit den anderen Aufgaben zwischen $r = .21$ und $.29$. Die Aufgaben 4 und 5 korrelierten miteinander am höchsten ($r = .41$). Auch die Aufgaben 2, 4 und 5 zeigten mit zwei bis vier anderen Aufgaben statistisch signifikante Zusammenhänge.

4.3 Korrelationen der offenen Aufgaben mit weiteren Leistungsmaßen

Nach der Z-Standardisierung aller Variablen wurden die Zusammenhänge der Leistungen in den offenen und geschlossenen Testaufgaben berechnet. Es zeigte sich, dass die Aufgabe 7 statistisch signifikant mit allen Leistungsmaßen korrelierte. Die Aufgaben 2 und 6 zeigten ebenfalls relativ viele statistisch signifikante Korrelationen mit den geschlossenen Aufgaben. Für die Aufgaben 1, 4 und 5 fand sich eine geringere Anzahl statistisch signifikanter Assoziationen. Wieder fällt die Aufgabe 3 mit niedrigen statistisch signifikanten Korrelationen ($.15 \leq r \leq .18$) auf (vgl. Tabelle 4).

⁸ Bei den Aufgaben 1 und 7 wurden öfter 0 Punkte vergeben, weil entweder nur Informationsfetzen oder pädagogische, aber keine historischen Argumente (z. B. »sehr interessant« ID 109202, »eindrucksvoll« ID 109210) genannt wurden.

Tabelle 4: Korrelationen mit weiteren Leistungsmaßen (*Korrelation ist auf dem Niveau von 0.05 zweiseitig signifikant. ** Korrelation ist auf dem Niveau von 0.01 zweiseitig signifikant.)

	Aufg. 1: De-Ko Verg.	Aufg. 2: De-Ko Gesch.	Aufg. 3: De-Ko Geg./ Zuk.	Aufg. 4: Re-Ko Verg. (A)	Aufg. 5: Re-Ko Verg. (ZZ)	Aufg. 6: Re-Ko Gesch.	Aufg. 7: Re-Ko Geg./ Zuk.
Sachkompetenz »Zeitzeuge«	.03	.23*	.13	.19*	.03	.11	.22**
Sachkompetenz »De-Ko«	.05	.27*	.18*	-.01	-.00	.14	.20**
Sachkompetenz »Re-Ko«	.13	.22*	.15*	.06	.12	.20**	.25**
Sachkompetenz »Q oder D?«	.18**	.27**	.04	.12	.25**	.28**	.31**

4.4 Exploratorische Faktorenanalysen

In der nachfolgenden exploratorischen Faktorenanalyse mit den sieben offenen Aufgaben ergab sich ein Ladungsmuster mit drei Faktoren, wobei der dritte Faktor die Aufgabe 3 repräsentierte. Eine zweite Faktorenanalyse ohne die Aufgabe 3 ergab zwei Faktoren, wobei die Aufgaben 4 und 5 auf dem ersten, und die Aufgaben 2, 6 und 7 auf dem zweiten Faktor lud. Die Aufgabe 1 wurde trotz einer Doppelladung dem Faktor 1 zugeordnet, da die Schülerinnen und Schüler in den Aufgaben 1, 4 und 5 in der Identifizierung von Vergangenheitspartikeln »Vergangenes erkennen« (Faktor 1). In den Aufgaben 2, 6 und 7 hingegen gingen die Schülerinnen und Schüler selbstständig mit Geschichte um, sodass der Faktor 2 mit »Geschichte denken« bezeichnet wurde (vgl. Tabelle 5).

Tabelle 5: Exploratorische Faktorenanalyse ohne Aufgabe 3. Hauptkomponentenanalyse: Promax-Rotation mit Kaiser-Normalisierung

	Faktor 1 (»Vergangenes erkennen«)	Faktor 2 (»Geschichte denken«)
Aufgabe 1 (De-Ko Verg.)	.49	.49
Aufgabe 2 (De-Ko Gesch.)	.04	.68
Aufgabe 4 (Re-Ko Verg./Aufruf)	.78	-.04
Aufgabe 5 (Re-Ko Verg./ZZ)	.82	.29
Aufgabe 6 (Re-Ko Gesch.)	.37	.55
Aufgabe 7 (Re-Ko Geg./Zuk.)	.07	.68

Die Faktoren korrelierten statistisch signifikant mit $r = .24$. Auf Basis der Faktorenanalysen wurden die Aufgaben durch eine Mittelwertbildung zu zwei Skalen zusammengefasst, wobei die Skala »Vergangenes erkennen« ein Cronbachs Alpha von .57 und die Skala »Geschichte denken« eines von .45 aufwies. Die statistisch signifikante Korrelation der beiden Skalen betrug $r = .28$.

4.5 Korrelationen der Skalen (offene und geschlossene Aufgaben)

Die Skala »Geschichte denken« korrelierte statistisch signifikant mit allen geschlossenen erfassten Kompetenztests, wohingegen die Skala »Vergangenheit erkennen« weniger und niedrigere Korrelationen mit diesen aufwies (vgl. Tabelle 6).

Tabelle 6: Korrelationen der Skalen der offenen Aufgaben mit geschlossen erfassten Kompetenzskalen (*Korrelation ist auf dem Niveau von 0.05 zweiseitig signifikant. ** Korrelation ist auf dem Niveau von 0.01 zweiseitig signifikant.)

	Vergangenes erkennen (Mittelwert der Aufgaben 1 + 4 + 5)	Geschichte denken (Mittelwert der Aufgaben 2 + 6 + 7)
Skala »SK Zeitzeuge«	.11	.20*
Skala »SK De-Ko«	.03	.26**
Skala »SK Re-Ko«	.17*	.29**
Skala »SK Q oder D?«	.27**	.38**

5. Zusammenfassung und Diskussion

Ziel der Untersuchung war es, auf Basis eines theoretisch begründeten Kompetenzmodells ein Messinstrument mit offenen Aufgabenformaten zu konzipieren und empirisch zu validieren. In den Befunden zeigen sich die Potenziale und Grenzen dieses Messinstruments.

Die Korrelation der auf Grundlage der Faktorenanalyse gebildeten beiden Skalen ist mit $r = .28$ eher niedrig. Auch wenn verschiedene Dimensionen der historischen Kompetenz gemessen werden, würde man doch eine höhere Korrelation zwischen den Teilbereichen erwarten. Allerdings kann man davon ausgehen, dass der »wahre« Wert der Korrelation der Skalen unterschätzt wird, da die Skalenwerte messfehlerbehaftet sind, was sich auch in der nicht zufriedenstellenden internen Konsistenz der beiden Skalen zeigt. Vor allem die geringe Item-Anzahl pro Skala (jeweils drei Aufgaben/Items) ist als Ursache für die Messungenauigkeit der Skalen in Betracht zu ziehen.

Auch wenn die Ergebnisse der Faktorenanalyse mit Vorsicht interpretiert werden müssen, da die Aufgaben (mit Ausnahme der Aufgaben 4 und 5) jeweils einen Teilbereich der Sechser-Matrix repräsentieren, werden doch geschichtstheoretisch relevante Einblicke eröffnet. Die beiden Faktoren »Vergangenes erkennen« und »Geschichte denken« scheinen – in einfacherer Form – die vertikalen Achsen in der Sechser-Matrix abzubilden: Fokussierung auf »Vergangenes« versus Fokussierungen auf »Geschichte« und »Gegenwart und Zukunft«. Dass die Aufgabe 3 auf keinem der beiden Faktoren lud, lässt sich aus dem Material erklären. Aus der sehr kompakten und informativen Darstellung des Schulbuchtexts die transportierte Botschaft herauszufiltern, stellte für Schülerinnen und Schüler der 9. Jahrgangsstufe vermutlich eine Überforderung dar.

Die vorliegenden Befunde deuten zum einen darauf hin, dass die vertikalen Achsen der Sechser-Matrix (= Fokussierungen) ähnliche Denkopoperationen abbilden. Dass das Herausfinden von Vergangenheitspartikeln aus einer Narration (Aufgabe 1) oder aus Quellen (Aufgaben 4 und 5) ähnliche Operationen verlangt, liegt auf der Hand. Doch auch analytische und interpretatorische Operationen scheinen im de- und rekonstruierenden Zugriff ähnliche Denkleistungen zu verlangen.

Die Abbildung der vertikalen Achsen der Sechser-Matrix in den beiden Faktoren könnte ebenso wie auch der Befund, dass im Faktor »Geschichte denken« die zwei Fokussierungen »Geschichte« und »Gegenwart und Zukunft« zusammenfallen, damit zusammenhängen, dass in den Aufgaben des zweiten Faktors

anspruchsvollere historische Denkopoperationen verlangt werden. Gerade diese Aufgaben weisen auch relativ hohe statistisch signifikante Korrelationen mit den geschlossenen Testaufgaben zu den epistemologischen Einsichten und dem Begriffsverständnis auf. Dies kann als Validitätshinweis sowohl hinsichtlich der Codierung und Skalierung der offenen Aufgaben als auch umgekehrt im Hinblick auf die geschlossenen Aufgabenformate verstanden werden.

In weiteren empirischen Studien wäre es lohnend, den Fragen nachzugehen, ob zum einen die Fokussierungen »Geschichte« und »Gegenwart/Zukunft« enger zusammengehören als theoretisch angenommen, und zum anderen, ob die idealtypische Unterscheidung von Re- und De-Konstruktionskompetenzen empirisch nachgewiesen werden kann. Die vorliegende Studie zeigt jedenfalls einen Weg, wie basierend auf dem theoretisch begründeten Modell historischen Denkens komplexe Kompetenzen in offenen Aufgaben operationalisiert und skaliert werden können.

Literatur

- Barricelli, Michele/Gautschi, Peter/Körper Andreas (2012): Historische Kompetenzen und Kompetenzmodelle. In: Barricelli, Michele/Lücke, Martin (Hrsg.): Handbuch Praxis des Geschichtsunterrichts (Band 1). Schwalbach/Wochenschau, S. 207–235.
- Berelson, Bernard (1952): Content Analysis in Communication Research. New York/Free Press.
- Bertram, Christiane/Wagner, Wolfgang/Trautwein, Ulrich (2013): Chancen und Risiken von Zeitzeugenbefragungen – Entwicklung eines Messinstruments für eine Interventionsstudie. In: Hodel, Jan/Waldis, Monika/Ziegler Béatrice (Hrsg.): Forschungswerkstatt Geschichtsdidaktik 12 – Beiträge zur Tagung »geschichtsdidaktik empirisch 12«. Bern/hep verlag, S. 108–119.
- Bertram, Christiane/Wagner, Wolfgang/Trautwein, Ulrich (2014): Zeitzeugenbefragungen im Geschichtsunterricht: Entwicklung eines Kurzinstruments für die Wirksamkeitsmessung. In: Arand, Tobias/Seidenfuss, Manfred (Hrsg.): Neue Wege – neue Themen – neue Methoden? Ein Querschnitt aus der geschichtsdidaktischen Forschung des wissenschaftlichen Nachwuchses. Göttingen/Vandenhoeck & Ruprecht, S. 191–208.
- Borries, Bodo von (2009): Historische Bildung als Voraussetzung für politische Mündigkeit? Verfügbar unter: http://www.lpb-mv.de/cms2/LfpB_prod/LfpB/_downloads/Jahreskongress_LpB_M-V_2009_Bodo_von_Borries.pdf [16.10.2013].
- Danto, Arthur C. (1980): Analytische Philosophie der Geschichte. Frankfurt am Main/Suhrkamp.
- Frederking, Volker (Hrsg.) (2008): Schwer messbare Kompetenzen. Herausforderungen für die empirische Fachdidaktik. Baltmannsweiler/Schneider Verlag Hohengehren.
- Frederking, Volker/Roick, Thorsten/Steinhauer, Lydia (2010): Literarästhetische Urteilskompetenz – Forschungsansatz und Zwischenergebnisse. In: Bayrhuber, Horst et al. (Hrsg.): Empirische Fundierung in den Fachdidaktiken – Fachdidaktische Forschung (Band 1). Münster/Waxmann, S. 75–94.

- Jeismann, Karl-Ernst (1977): Didaktik der Geschichte – Die Wissenschaft von Zustand, Funktion und Veränderung geschichtlicher Vorstellungen im Selbstverständnis der Gegenwart. In: Kosthorst, Erich (Hrsg.): *Geschichtswissenschaft – Didaktik, Forschung, Theorie*. Göttingen/Vandenhoeck & Ruprecht, S. 9–33.
- Klieme, Eckhard et al. (2003): Zur Entwicklung nationaler Bildungsstandards – Eine Expertise. Verfügbar unter: http://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf [14.10.2013].
- Körber, Andreas/Schreiber, Waltraud/Schöner, Alexander (Hrsg.) (2007): *Kompetenzen historischen Denkens – Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Kompetenzen: Grundlagen – Entwicklung – Förderung*. Neuried/ars una.
- Kuckartz, Udo (2007): *Einführung in die computergestützte Analyse qualitativer Daten*. Wiesbaden/VS.
- Kühberger, Christoph (2009): *Kompetenzorientiertes historisches und politisches Lernen – Methodische und didaktische Annäherungen an Geschichte, Sozialkunde und politische Bildung*. Innsbruck/Studien Verlag.
- Mayring, Philipp (2010): *Qualitative Inhaltsanalyse – Grundlagen und Techniken*. (11. Auflage). Weinheim, Basel/Beltz.
- Pandel, Hans-Jürgen (2010): *Historisches Erzählen – Narrativität im Geschichtsunterricht*. Schwalbach/Wochenschau.
- Ricœur, Paul (1988): *Zeit und Erzählung*. München/Wilhelm Fink Verlag.
- Rüsen, Jörn (Hrsg.) (1983): *Historische Vernunft – Grundzüge einer Historik I: Die Grundlagen der Geschichtswissenschaft*. Göttingen/Vandenhoeck & Ruprecht.
- Rüsen, Jörn (2013): *Historik – Theorie der Geschichtswissenschaft*. Köln/Böhlau.
- Schöner, Alexander (2007): *Kompetenzbereich historische Sachkompetenzen*. In: Körber, Andreas/Schreiber, Waltraud/Schöner, Alexander (Hrsg.): *Kompetenzen historischen Denkens – Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Kompetenzen: Grundlagen – Entwicklung – Förderung*. Neuried/ars una, S. 265–314.
- Schreiber, Waltraud et al. (2007): *Historisches Denken – Ein Kompetenz-Strukturmodell (Basisbeitrag)*. In: Körber, Andreas/Schreiber, Waltraud/Schöner, Alexander (Hrsg.): *Kompetenzen historischen Denkens – Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Kompetenzen: Grundlagen – Entwicklung – Förderung*. Neuried/ars una, S. 17–53.
- Weinert, Franz E. (2001): *Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit*. In: Weinert, Franz E. (Hrsg.): *Leistungsmessungen in Schulen*. Weinheim, Basel/Beltz, S. 17–32.