

Technische Universität Dortmund
Fakultät für Erziehungswissenschaften, Psychologie und Bildungsforschung

Civic Literacy im Jugendalter
Konzeption, Operationalisierung und empirische Erfassung

Kumulative Dissertation zur Erlangung des akademischen Grades
Doktor der Philosophie (Dr. phil.)

Vorgelegt von
Pascal Ulrich Alscher, M. Sc.
Geboren am 22.03.1995 in Osnabrück
Matrikelnummer 223385

Erstgutachterin: Prof. Dr. Nele McElvany
Zweitgutachter: Prof. Dr. Olaf Köller

Dortmund, Januar 2024

Dissertation in der Fakultät für Erziehungswissenschaft, Psychologie und Bildungsforschung
an der Technischen Universität Dortmund

Danksagung

Die Zeit, die ich mit der Arbeit an dieser Dissertation verbracht habe, war eine der schönsten Zeiten meines Lebens, die ich nicht missen möchte. Eine Reihe von Menschen haben maßgeblich zum Gelingen dieser Arbeit beigetragen. Ihnen allen möchte ich an dieser Stelle meinen herzlichen Dank aussprechen.

Zunächst möchte ich Frau Prof. Dr. Nele McElvany für die stets konstruktive, ermutigende und immer von hoher Integrität geprägte Betreuung meiner Dissertation bedanken. Insbesondere möchte ich Ihr für die ausgezeichnete Balance zwischen Unterstützung und Anleitung einerseits sowie der Freiheit, eigene Ideen zu entwickeln und voranzutreiben, andererseits danken. Außerdem gilt mein Dank Herrn Prof. Dr. Olaf Köller, der die Arbeit als Zweitgutachter begleitet hat. Mein Dank gilt zudem Herrn Prof. Dr. Hermann Josef Abs, der die Arbeit zu unterschiedlichen Zeitpunkten mit seiner Expertise und stets konstruktiven Rückmeldungen unterstützt hat und zugleich als Drittgutachter der Arbeit fungierte.

Einen großen Dank möchte ich zudem Herrn Dr. Ulrich Ludewig aussprechen, der mich mit wertvollen Anregungen unterstützt hat. Dein offenes Ohr, dein Enthusiasmus für die Forschung und dein Optimismus waren insbesondere zu Beginn meiner Promotion von unschätzbarem Wert. Des Weiteren möchte ich allen Kolleginnen und Kollegen am Institut für Schulentwicklungsforschung für die überaus angenehme und von gegenseitiger Unterstützung geprägte Arbeitsatmosphäre danken; das macht Vieles leichter. Ich gehe gerne zur Arbeit und dafür seid vor allem Ihr verantwortlich.

Schließlich möchte ich mich bei meiner Familie und meinen Freundinnen und Freunden dafür bedanken, dass sie mir in schwierigen Zeiten zur Seite standen. Ohne die vielfältige Unterstützung meiner Familie wären viele Dinge nicht möglich gewesen. Ein besonderer Dank gilt zudem meiner Verlobten. Aizhana, your love, patience and understanding have given me the strength to overcome any challenges along the way. Your belief in me has always motivated me to keep going and do my best.

Der letzte Dank gebührt meiner Mutter, die mich immer dazu ermuntert hat, meinen Träumen nachzugehen und die mich in jeder Lebenslage unterstützt hat. Es schmerzt zutiefst, dass Du so plötzlich von uns gegangen bist. Dass ich deinen letzten Abend zusammen mit Dir bei meiner Disputation verbringen durfte, wird mir immer ein Trost sein.

Zusammenfassung

Bildungsstandards werden als wichtiges Instrument der Überprüfung und Steuerung des Bildungssystems angesehen (Köller, 2020). Anders als in anderen Schulfächern führte die zu Beginn der 2000er Jahre geführte Kompetenzdebatte in der deutschen Bildungspolitik jedoch zu keinen national verbindlichen Bildungsstandards für die politische Bildung. Stattdessen wurden unterschiedliche Versuche unternommen, ein Kompetenzmodell für die schulische politische Bildung zu entwickeln (z. B. Autorengruppe Fachdidaktik, 2011; Detjen et al., 2012; GPJE, 2004). Anspruch der entwickelten Kompetenzmodelle war vor allem zwischen abstrakten Bildungszielen und konkreten Aufgabenstellungen zu vermitteln und den kompetenzorientierten Unterricht an Schulen anzuleiten und zu erleichtern (Detjen et al., 2012; Klieme et al., 2003). Zudem sollten die Modelle als Orientierungsrahmen für empirische Studien dienen und die Möglichkeit schaffen, wissenschaftliche Fragestellungen zu entwickeln und zu prüfen (Detjen et al., 2012). Gleichwohl ist eine Operationalisierung der entwickelten Kompetenzmodelle nur für ausgewählte Teilbereiche gegeben.

Ausgehend von dieser Angebotslücke wurde im Rahmen dieser Dissertation ein Civic Literacy Kompetenzmodell entwickelt und empirisch überprüft. Die Dissertation umfasst einen konzeptuellen sowie vier empirische Beiträge. Konkret wurde untersucht, (1) was Civic Literacy umfasst und wie sie bei Jugendlichen gemessen werden kann, (2) inwiefern unterschiedliche Merkmale der schulischen politischen Bildung mit den verschiedenen Bereichen von Civic Literacy in Zusammenhang stehen und (3) wie subjektives und objektives politisches Wissen zusammenhängen. Für die Beantwortung der Fragen wurde auf die Sekundarstufe I und insbesondere die Klassenstufen 7 und 10 als Entwicklungsphase fokussiert.

In *Beitrag I* wurde Civic Literacy zunächst als das Wissen und die Fähigkeit, informierte Entscheidungen hinsichtlich gesellschaftlich relevanter Fragestellungen zu treffen, und die für die Umsetzung der getroffenen Entscheidungen prinzipiell notwendigen Handlungsdispositionen (d. h. Motivationen, Einstellungen und Volitionen) zur politischen und zivilgesellschaftlichen Partizipation definiert (z. B. Kultusministerkonferenz 2018; Milner 2002; Partnership for 21st Century Skills 2009). Aufbauend auf dieser Definition sowie theoretischen Vorüberlungen (z. B. Autorengruppe Fachdidaktik, 2011, Detjen et al., 2012; Weißeno et al., 2010) und empirischen Untersuchungen (z. B. Schulz et al., 2018) wurde ein Civic Literacy Kompetenzmodell für die Sekundarstufe I konzeptualisiert und in Teilen operationalisiert. Für das Kompetenzmodell wurden die Dimensionen politisches Wissen, politische

Motivationen, politische Einstellungen und politische Volitionen unterschieden. Für die Dimension des politischen Wissens wurde ein Wissenstest speziell für Schüler*innen der Sekundarstufe I entwickelt. Die weiteren Dimensionen wurden in die Facetten politisches Interesse und politisches Effektivitätsgefühl (politische Motivationen), Einstellung zur Demokratie beziehungsweise zur pluralen Gesellschaft (politische Einstellungen) und politische beziehungsweise zivilgesellschaftliche Partizipationsbereitschaft (politische Volitionen) unterteilt.

In *Beitrag II* wurde zunächst die in Deutschland geführte Debatte zur Einführung von national verbindlichen Bildungsstandards thematisiert (Köller, 2016) und auf deren Bedeutung für die deutsche Politikdidaktik fokussiert (z. B. Hedinger, 2018; Massing, 2022). Anschließend wurden die theoretischen Annahmen des entwickelten Civic Literacy Kompetenzmodells anhand empirischer Daten überprüft. Hierfür wurden die Daten von insgesamt $N = 1.047$ Schüler*innen der 7. ($n = 613$) und 10. Klassenstufe ($n = 434$) genutzt. Konkret wurde die theoretisch angenommene Binnenstruktur des Modells, die Messinvarianz in den beiden Klassenstufen, die Entwicklungs- beziehungsweise Alterssensibilität sowie die Konvergenz und Diskriminanz der Modellinstrumente überprüft. Die theoretisch angenommene Struktur wurde weitgehend durch konfirmatorische Faktorenanalysen gestützt. Ausnahmen betrafen das politische Effektivitätsgefühl und die politische beziehungsweise zivilgesellschaftliche Partizipationsbereitschaft. Für das politische Effektivitätsgefühl empfiehlt sich eine Unterscheidung in internes und externes politisches Effektivitätsgefühl (Oberle & Wenzel, 2019). Die Unterscheidung der beiden Facetten politische und zivilgesellschaftliche Partizipationsbereitschaft als Teil der Dimension politische Volition wurde nicht durch die Daten gestützt (siehe jedoch Ekman & Amnå, 2012). Mehr-Gruppen konfirmatorische Faktorenanalysen zeigten, dass skalare Messinvarianz für den Einsatz der Modellinstrumente in der 7. und 10. Klassenstufe angenommen werden kann. Bezüglich der Alterssensibilität zeigte sich weiter, dass die Mittelwerte aller Facetten des Kompetenzmodells in der 10. Klassenstufe höher waren als in der 7. Klassenstufe. Auch die Ergebnisse zur Konvergenz und Diskriminanz der Modellinstrumente stützten die zuvor im Rahmen eines nomologischen Netzwerks gemachten theoretischen Annahmen weitgehend. Die theoretischen Modellannahmen wurden fast vollständig von den empirischen Daten gestützt. Die Ergebnisse legen nahe, dass sich das Modell und die Instrumente für den Einsatz in längsschnittlichen Studien in der Sekundarstufe I eignen.

In *Beitrag III* wurden der politische Wissenstest und die Schwierigkeit der Testitems in den Fokus genommen. Idealerweise wird die Schwierigkeit eines Items ausschließlich durch die Aufgabenanforderung in Bezug auf die zu messende Merkmalsausprägung bestimmt (Embretson & Reise, 2013). In der Realität gibt es jedoch eine Vielzahl weiterer Faktoren, die die

Schwierigkeit von Testitems beeinflussen. Neben externen Störfaktoren wie Testmotivation und Testängstlichkeit (z. B. Spoden et al., 2020) spielt auch der Expositionsgrad der Iteminhalte eine Rolle. Ein vielfach genutzter Proxy für den Expositionsgrad von Iteminhalten ist Worthäufigkeit (Brysbaert et al., 2018). Insbesondere im Kontext politischer Bildung als prinzipiell sprachintensive Lerndomäne (Drippe, 2018), ist davon auszugehen, dass Itemschwierigkeiten durch durchschnittliche Worthäufigkeiten beeinflusst werden könnten. Vor dem Hintergrund der stark ausgeprägten Fachsprache in der politischen Bildung (GPJE, 2004) kann zudem eine Unterscheidung in Worthäufigkeiten in alltäglichen und akademischen Sprachsettings relevant sein. Konkret wurde untersucht, inwiefern die durchschnittliche Häufigkeit von Wörtern eines Items in alltäglichen und akademischen Sprachsettings mit der Itemschwierigkeit zusammenhängen. Für die Untersuchung wurden Daten von insgesamt $N = 250$ Schüler*innen der 7. ($n = 152$) und 10. Klassenstufe ($n = 98$) genutzt. Die Ergebnisse zeigten, dass eine hohe Worthäufigkeit in alltäglichen Sprachsettings negativ mit der Itemschwierigkeit zusammenhängt. Zwischen der Worthäufigkeit in akademischen Sprachsettings und der Itemschwierigkeit bestand kein Zusammenhang. Wenn beide Indikatoren der Worthäufigkeit in einem gemeinsamen Modell untersucht wurden, hing die alltägliche Worthäufigkeit weiter statistisch signifikant negativ mit der Itemschwierigkeit zusammen. Zusätzlich hing die akademische Worthäufigkeit positiv mit der Itemschwierigkeit zusammen. Die Forschungsbefunde verdeutlichen die Bedeutung von Worthäufigkeiten für die Schwierigkeit von Items in politischen Wissenstests. Insbesondere relative Unterschiede zwischen Worthäufigkeiten in alltäglichen und akademischen Sprachsettings haben das Potential, Varianz in der Itemschwierigkeit aufzuklären.

In *Beitrag IV* wurde der Zusammenhang zwischen unterschiedlichen Merkmalen der Unterrichtsqualität politischer Bildung und verschiedenen Bereichen des Civic Literacy Kompetenzmodells untersucht. Es gilt als weithin gesichert, dass Unterrichtsqualität eine wichtige Rolle für das Erreichen von Lernzielen spielt (z. B. Praetorius et al., 2017). Im Kontext politischer Bildung sind fachspezifische Merkmale einer Unterrichtsqualität bislang jedoch nur in Teilen erforscht (siehe jedoch Weißeno & Grobshäuser, 2021). Einen Anhaltspunkt für ein fachspezifisches Modell der Unterrichtsqualität bilden insbesondere das offene Unterrichtsklima (z. B. Eckstein et al., 2021), die kognitive Aktivierung (Weißeno & Landwehr, 2015) und das Diskutieren aktueller politischer Themen (Bittman & Russel, 2016). Gerade in Bezug auf die relative Bedeutung verschiedener Merkmale der Unterrichtsqualität liegen für die unterschiedlichen Zielvorgaben der politischen Bildung bislang kaum empirische Befunde vor (siehe jedoch Bayram-Özdemir et al., 2016). Auf Basis der Daten von insgesamt $N = 250$ Schüler*innen der 7. ($n = 152$) und 10. ($n = 98$) Klassenstufe wurde untersucht, inwiefern ein als offen

wahrgenommenes Unterrichtsklima, ein als kognitiv aktivierend wahrgenommener Unterricht sowie das Diskutieren aktueller politischer Themen mit der Partizipationsbereitschaft von Schüler*innen zusammenhängt und wie diese Zusammenhänge durch politisches Interesse und politisches Wissen mediiert werden. Zur Beantwortung der Forschungsfragen wurden Strukturgleichungsmodelle gerechnet. Das offene Unterrichtsklima hing direkt und positiv mit der Partizipationsbereitschaft zusammen. Allerdings bestand ein negativer Zusammenhang zwischen dem offenen Unterrichtsklima und dem politischen Wissen. Der Zusammenhang zwischen der kognitiven Aktivierung und der Partizipationsbereitschaft wurde vollständig über das politische Interesse und das politische Wissen mediiert. Das Diskutieren aktueller politischer Themen hing positiv mit dem politischen Interesse zusammen. Eine separate Analyse der Daten der 7. und 10. Klassenstufe bestätigte das Befundmuster weitgehend. Die Ergebnisse verdeutlichen die Wichtigkeit der Unterrichtsqualität im Kontext politischer Bildung und unterstreichen die Bedeutung eines zieldifferenten Einsatzes unterschiedlicher Unterrichtsmethoden.

In *Beitrag V* wird der Dunning-Kruger Effekt in Bezug auf das politische Wissen untersucht. Der Dunning-Kruger Effekt beschreibt das Befundmuster, dass leistungsschwache Individuen ihre eigene Fähigkeit oder Testleistung weniger genau einschätzen als leistungsstarke Individuen (Kruger & Dunning, 1999). Zudem neigen leistungsschwache Individuen eher zu einer Überschätzung und leistungsstarke Individuen eher zu einer Unterschätzung. Die klassische Interpretation des Dunning-Kruger Effekts geht davon aus, dass leistungsschwache Individuen nicht die metakognitiven Fähigkeiten haben, um ihr Leistungs- beziehungsweise Fähigkeitsdefizit erkennen zu können (z. B. Dunning, 2011; Schlösser et al., 2013). Kritiker*innen hingegen nehmen an, dass der Dunning-Kruger Effekt ein statistisches Artefakt ist (z. B. Gignac & Zajenkowski, 2020; McIntosh et al., 2022). Insbesondere im Schulkontext spielen Selbsteinschätzungen und die Fähigkeiten, eigene Lernfortschritte beurteilen und Lerndefizite erkennen zu können, eine große Rolle (z. B. Ohtani & Hisasaka, 2018; Perry et al., 2019). Für gewöhnlich sind diese Fähigkeiten in höheren Klassenstufen stärker ausgeprägt, als in niedrigeren Klassenstufen (z. B. Artelt et al., 2012; Van der Stel & Veenman, 2010). Es wurde untersucht, ob der Dunning-Kruger Effekt in Bezug auf das politische Wissen bei Schüler*innen der Sekundarstufe I existiert und inwiefern es Unterschiede zwischen Schüler*innen der 7. und 10. Klassenstufe gibt. Zudem wurde untersucht, wie die wahrgenommene Motivierungsqualität und kognitive Aktivierung mit der Genauigkeit und Richtung der Selbsteinschätzungen zusammenhängt. Neben der klassischen visuellen Überprüfung wurden statistische Verfahren

eingesetzt, um den Dunning-Kruger Effekt zu untersuchen. Außerdem wurden Strukturgleichungsmodelle mit Urteilsgenauigkeit und -richtung als abhängige und Motivierungsqualität und kognitive Aktivierung als unabhängige Variable gerechnet. Für die Analysen wurden die Daten von insgesamt $N = 1.047$ Schüler*innen der 7. ($n = 613$) und 10. Klassenstufe ($n = 434$) genutzt. Sowohl die visuelle Überprüfung als auch die statistischen Analysen zeigten, dass das für den Dunning-Kruger Effekt typische Befundmuster in beiden Klassenstufen in ähnlicher Form existierte. Zudem hing die wahrgenommene Motivierungsqualität positiv mit der Urteilsgenauigkeit, aber auch mit der Wahrscheinlichkeit einer Überschätzung zusammen. Die kognitive Aktivierung hing hingegen negativ mit der Urteilsgenauigkeit zusammen. Die Beständigkeit mit der sich der Effekt in zwei aus einer Entwicklungsperspektive so unterschiedlichen Gruppe von Individuen zeigte, könnte ein Indiz dafür sein, dass der Effekt tatsächlich ein statistisches Artefakt ist.

In der studienübergreifenden Diskussion werden die Ergebnisse der Beiträge aufeinander bezogen. Die Beiträge I, II und III zeigen, dass ein funktionsfähiges Civic Literacy Kompetenzmodell mit einem eigens für Schüler*innen der Sekundarstufe I konzipierten Wissenstest sowie teilweise neuentwickelten Skalen für die weiteren Modellbereiche ausgearbeitet wurde. Es werden zudem mögliche Anwendungsfelder außerhalb der Sekundarstufe I und des deutschen Bildungssystems diskutiert (Holle & Ványi, 2022; Marci-Boehncke et al., 2023). Beitrag IV zeigt die Bedeutung eines zieldifferenten Einsatzes unterschiedlicher Unterrichtsmethoden im Kontext politischer Bildung auf. Es wird diskutiert, inwiefern die Ergebnisse des Beitrags zur Theoriebildung für die Tiefenstrukturen qualitativ hochwertigen Politikunterrichts beitragen. Zudem wird Bezug genommen auf weitere in diesem Kontext durchgeführte Studien, die beispielsweise auf Emotionen (Graf et al., 2023) oder die sozialen Dynamiken im Kontext politischer Bildung fokussieren (Stattin & Russo, 2023). Die Ergebnisse von Beitrag V verdeutlichen, dass viele Schüler*innen der Sekundarstufe I ihr politisches Wissen systematisch falsch einschätzen und unterstreichen die Notwendigkeit einer objektiven Erfassung.

Diese Dissertation erweitert die existierende Forschung im Bereich der politischen Bildung von Schüler*innen der Sekundarstufe I, indem ein vollständig operationalisiertes und empirisch überprüftes Civic Literacy Kompetenzmodell entwickelt wurde. Die Ergebnisse geben Aufschluss über das Niveau der unterschiedlichen Kompetenzbereiche bei Schüler*innen der Klassenstufe 7 und 10. Zudem wird untersucht, wie die Modellbereiche untereinander aber auch mit weiteren für den Bildungskontext relevanten Variablen zusammenhängen. Durch den Einsatz des Kompetenzmodells in zukünftigen Forschungsprojekten kann ein wichtiger Beitrag für die Beforschung und Praxis politischer Bildung geleistet werden.

Abstract

Educational standards are seen as an important instrument for reviewing and steering the education system (Köller, 2020). Unlike in other school subjects, however, the competence debate in German education policy at the beginning of the 2000s did not lead to any nationally binding educational standards for civic education. Instead, various attempts were made to develop a competence model for civic education in schools (e.g., Autorengruppe Fachdidaktik, 2011; Detjen et al., 2012; GPJE, 2004). The main goal of the developed competence models was to mediate between abstract educational goals and concrete tasks as well as to guide and facilitate competence-oriented teaching in schools (Detjen et al., 2012; Klieme et al., 2003). In addition, the models should serve as a framework for empirical studies and create the opportunity to develop and test scientific questions (Detjen et al., 2012). Nevertheless, the so far developed competence models have only been operationalized for selected sub-areas.

Based on this gap, a civic literacy competence model was developed and empirically tested as part of this dissertation. The dissertation comprises one conceptual and four empirical contributions. Specifically, it examines (1) what civic literacy comprises and how it can be measured in youth, (2) the extent to which different characteristics of civic education at school are related to the various areas of the developed civic literacy framework and (3) how subjective and objective political knowledge are related. To answer these questions, we focused on students in secondary education and in particular students in grades 7 and 10 as a developmental phase.

In *Study I*, civic literacy was initially defined as the knowledge and ability to make informed decisions regarding politically relevant issues and the dispositions (i.e., motivations, attitudes and volitions) for political and civic participation that are fundamentally necessary for the implementation of the decisions made (e.g., Kultusministerkonferenz 2018; Milner 2002; Partnership for 21st Century Skills 2009). Based on this definition as well as theoretical consideration (e.g., Autorengruppe Fachdidaktik, 2011, Detjen et al., 2012; Wei-ßenö et al., 2010) and empirical evidence (e.g., Schulz et al., 2018), a civic literacy competence model for high school students was conceptualized and partially operationalized. For the competence model, a distinction was made between the dimensions of political knowledge, political motivations, political attitudes and political volitions. For the dimension of political knowledge, a knowledge test was developed specifically for high school students. The other dimensions were divided into the facets of political interest and sense of political efficacy (political motivations), attitudes towards democracy and a pluralistic society (political attitudes) and willingness to politically and civically participate (political volitions).

Study II started with addressing the debate in Germany on the introduction of nationally binding educational standards (Köller, 2016) and its significance for German political didactics (e.g., Hedinger, 2018; Massing, 2022). The theoretical assumptions of the developed civic literacy competence model were then tested using empirical data. The data from a total of $N = 1,047$ students in grades 7 ($n = 613$) and 10 ($n = 434$) were used for this purpose. Specifically, the theoretically assumed internal structure of the model, the measurement invariance in the two grades, the age sensitivity as well as the convergent and discriminant validity of the model instruments were examined. The theoretically assumed structure was largely supported by confirmatory factor analyses. Exceptions concerned the sense of political efficacy and the willingness to politically and civically participate. For the sense of political efficacy a distinction between internal and external political efficacy is recommended (Oberle & Wenzel, 2019). The distinction between the two facets of political and civic willingness to participate as part of the political volition dimension was not supported by the data (see, however, Ekman & Amnå, 2012). Multi-group confirmatory factor analyses showed that scalar measurement invariance can be assumed for the use of the model instruments in the 7th and 10th grade. With regard to age sensitivity, it was also shown that the mean values of all facets of the competence model were higher in the 10th grade than in the 7th grade. The results on the convergence and discriminance of the model instruments also largely supported the theoretical assumptions previously made in the context of a nomological network. The theoretical model assumptions were almost completely supported by the empirical data. The results suggest that the model and the instruments are suitable for use in longitudinal studies in lower secondary education.

Study III focused on the political knowledge test and the difficulty of the test items. Ideally, the difficulty of an item is determined solely by the task requirement in relation to the attribute to be measured (Embretson & Reise, 2013). In reality, however, there are many other factors that influence the difficulty of test items. In addition to external confounding factors such as test motivation and test anxiety (e.g. Spoden et al., 2020), the degree of exposure of the item content also plays a role. A frequently used proxy for the degree of exposure of item content is word frequency (Brysbaert et al., 2018). Particularly in the context of civic education as a language-intensive learning domain (Drippe, 2018), it can be assumed that item difficulty could be influenced by average word frequencies. Against the background of the strongly pronounced technical language in civic education (GPJE, 2004), a distinction between word frequencies in everyday and academic language settings may also be relevant. Specifically, it was investigated to what extent an item's average word frequency in everyday and academic language settings was related to item difficulty. Data from a total of $N = 250$ students in the 7th

($n = 152$) and 10th grade ($n = 98$) were used for the study. The results showed that a high word frequency in everyday language settings was negatively related to item difficulty. There was no correlation between word frequency in academic language settings and item difficulty. When both indicators of word frequency were examined in a joint model, everyday word frequency continued to be statistically significantly negatively related to item difficulty. In addition, academic word frequency was positively related to item difficulty. The research findings illustrate the importance of word frequency for the difficulty of items in political knowledge tests. In particular, relative differences between word frequencies in everyday and academic language settings have the potential to explain variance in item difficulty.

In *Study IV*, the relationship between different characteristics of the quality of civic education teaching and various areas of the civic literacy competence model was examined. It is widely accepted that teaching quality plays an important role in the achievement of learning objectives (e.g. Praetorius et al., 2017). In the context of civic education, however, subject-specific characteristics of teaching quality have so far only been partially researched (see, however, Weißeno & Grobshäuser, 2021). The open classroom climate (e.g. Eckstein et al., 2021), cognitive activation (Weißeno & Landwehr, 2015) and the discussion of current political topics (Bittman & Russel, 2016) are particularly indicative of a subject-specific model of teaching quality. However, there is so far only little empirical evidence with regard to the relative importance of various characteristics of teaching quality for the different objectives of civic education (see however Bayram-Özdemir et al., 2016). Based on data from a total of $N = 250$ students in grades 7 ($n = 152$) and 10 ($n = 98$), we investigated the extent to which a classroom climate perceived as open, lessons perceived as cognitively activating and the discussion of current political topics are related to students' willingness to participate and how these relationships are mediated by political interest and political knowledge. Structural equation models were calculated to answer the research questions. The open classroom climate was directly and positively related to willingness to participate. However, there was a negative correlation between open classroom climate and political knowledge. The correlation between cognitive activation and willingness to participate was mediated entirely by political interest and political knowledge. Discussing current political issues was positively related to political interest. A separate analysis of the data from the 7th and 10th grade largely confirmed the pattern of findings. The results illustrate the importance of teaching quality in the context of civic education and underline the importance of using different teaching methods for different purposes.

In *Study V*, we examined the Dunning-Kruger effect in relation to political knowledge. The Dunning-Kruger effect describes the pattern of findings that low-achieving individuals assess their own ability or test performance less accurately than high-achieving individuals (Kruger & Dunning, 1999). In addition, low-performing individuals tend to overestimate and high-performing individuals tend to underestimate their ability or performance. The classic interpretation of the Dunning-Kruger effect assumes that low-performing individuals do not have the metacognitive skills to recognize their ability or performance deficit (e.g., Dunning, 2011; Schlösser et al., 2013). Critics, on the other hand, assume that the Dunning-Kruger effect is a statistical artifact (e.g., Gignac & Zajenkowski, 2020; McIntosh et al., 2022). Self-assessment and the ability to assess one's own learning progress and recognize learning deficits play a particularly important role in the school context (e.g., Ohtani & Hisasaka, 2018; Perry et al., 2019). These skills are usually more pronounced in higher grades than in lower grades (e.g., Artelt et al., 2012; Van der Stel & Veenman, 2010). It was thus investigated whether the Dunning-Kruger effect exists in relation to political knowledge in high school students and to what extent there are differences between students in grades 7 and 10. In addition, it was investigated how the perceived quality of motivation and cognitive activation in class is related to the accuracy and direction of the self-assessments. In addition to the classic visual inspection, statistical methods were used to investigate the Dunning-Kruger effect. Furthermore, structural equation models were calculated with judgment accuracy and judgement direction as dependent variables and motivational quality and cognitive activation as independent variables. The data from a total of $N = 1,047$ students in grades 7 ($n = 613$) and 10 ($n = 434$) were used for the analyses. Both the visual inspection and the statistical analyses showed that the pattern of findings typical of the Dunning-Kruger effect existed in a similar form in both grades. In addition, the perceived motivational quality was positively related to judgment accuracy, but also to the probability of overestimation. Cognitive activation, was negatively related to judgment accuracy. The consistency with which the effect was shown in two groups of individuals who were so different from a developmental perspective could be an indication that the effect is actually a statistical artifact.

In the general discussion, the results of the articles are related to each other. Studies I, II and III show that a functional civic literacy competence model was developed with a knowledge test specifically designed for high school school students and partially newly developed scales for the other model areas. Possible fields of application outside of the German secondary school system are discussed (Holle & Ványi, 2022; Marci-Boehncke et al., 2023). Study IV highlights the importance of using different teaching methods in the context of civic

education. It discusses the extent to which the results of the analyses contribute to the development of theory for the in-depth structures of high-quality civic education. Reference is also made to other studies conducted in this context, which focus, for example, on emotions (Graf et al., 2023) or social dynamics in the context of civic education (Stattin & Russo, 2023). The findings of Study V illustrate that many high school students systematically misperceive their political knowledge and underscore the need for an objective assessment political knowledge.

This dissertation extends existing research in the field of civic literacy among high school students by developing a fully operationalized and empirically validated civic literacy competence model. The results provide information about the level of the different competence areas among students in grades 7 and 10 and examine how the model areas relate to each other as well as to other variables relevant to the educational context. By using the competence model in future research projects, an important contribution can be made to the research and practice of civic education.

Inhaltsverzeichnis

Zusammenfassung	I
Abstract	VI
Tabellenverzeichnis.....	XIV
Abbildungsverzeichnis	XVI
1 Einleitung.....	1
1.1 Demokratie und Bildung	3
1.2 Politische Bildung in Deutschland und im internationalen Kontext	5
1.2.1 Ausgangspunkt und Entwicklung.....	7
1.2.2 Relevante Akteur*innen	10
1.3 Politische Bildung in der Sekundarstufe I.....	12
1.3.1 Bezeichnung und Umfang	13
1.3.2 Unterrichtsqualität	14
1.3.3 Fachfremder Unterricht	15
1.4 Bildungsstandards und Kompetenzmodelle für die politische Bildung	17
1.4.1 Kompetenzmodelle für die politische Bildung in Deutschland.....	18
1.4.2 Standards und Assessment Frameworks politischer Bildung im internationalen Kontext	20
1.5 Herausforderungen für die empirische Erfassung von Kompetenzen.....	21
1.5.1 Qualitätsanforderungen an psychologische Tests.....	23
1.5.2 Herausforderungen bei der Überprüfung und Interpretation von Testwerten	25
1.5.3 Herausforderungen im Kontext der empirischen Erfassung von Civic Literacy.....	27
1.6 Forschungsfragen	31
1.7 Zusammenfassungen der Einzelbeiträge.....	33
1.7.1 Beitrag I: Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung	33

1.7.2	Beitrag II: Towards the Measurability of a Competence Model for School-based Civic Education.....	34
1.7.3	Beitrag III: What Makes Domain Knowledge Difficult? Word Usage Frequency from SUBTLEX and dlexDB Explains Knowledge Item Difficulty	35
1.7.4	Beitrag IV: Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators	37
1.7.5	Beitrag V (weiterführende Analysen): When Will They Know What They Don't Know? – Political knowledge and the Infamous Dunning-Kruger Effect	38
	Literaturverzeichnis I.....	42
2	Beiträge der kumulativen Dissertation.....	63
2.1	Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung.....	63
2.2	Towards the Measurability of a Competence Model for School-based Civic Education.....	90
2.3	What Makes Domain Knowledge Difficult? Word Usage Frequency from SUBTLEX and dlexDB Explains Knowledge Item Difficulty	127
2.4	Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators	161
3	Weiterführende Analysen	191
3.1	When Will They Know What They Don't Know? – Political knowledge and the Infamous Dunning-Kruger Effect	191
4	Diskussion.....	226
4.1	Diskussion zentraler Ergebnisse.....	226
4.1.1	Was umfasst Civic Literacy und wie kann sie bei Jugendlichen gemessen werden?.....	226
4.1.2	Inwiefern stehen unterschiedliche Merkmale der schulischen politischen Bildung mit den verschiedenen Bereichen von Civic Literacy in Zusammenhang?.....	231

4.1.3	Wie hängen subjektives und objektives politisches Wissen zusammen??	233
4.2	Stärken und Limitation	236
4.2.1	Stärken	236
4.2.2	Limitationen	238
4.3	Implikationen für Forschung und Praxis	240
4.3.1	Implikationen für die Forschung	240
4.3.2	Implikationen für die Praxis	242
4.4	Fazit	244
	Literaturverzeichnis II	245

Tabellenverzeichnis

Einleitung

Tabelle 1a.	<i>Veröffentlichte Beiträge der kumulativen Dissertation</i>	41
Tabelle 1b.	<i>Beiträge zu weiterführenden Analysen</i>	41

Beitrag II

Table 1.	<i>Means, standard deviations, and correlations</i>	106
Table 2.	<i>Comparison of Fit Indices in different nested CFAs</i>	107
Table 3.	<i>Measurement invariance between 7th and 10th grade</i>	109

Anhang Beitrag II

Table A1.	<i>Means of framework constructs in grade 7 and 10</i>	122
Table A2.	<i>Correlations within and across framework dimensions and with sense of social affiliation</i>	123
Table B1.	<i>Comparison of Fit Indices in different CFAs</i>	126

Beitrag III

Table 1.	<i>Top 30 most frequent (see Figure 2; triangles) words (nouns, verbs, and adjectives) and most infrequent words (circles)</i>	140
Table 2.	<i>Top 30 words (nouns, verbs, and adjectives) more frequent in everyday settings (see Figure 2; crosses) and more frequent in academic settings (squares)</i>	141
Table 3.	<i>Correlations and descriptive statistics</i>	142
Table 4.	<i>Regression analysis explaining item difficulty in the political knowledge test.</i>	143
Table 5.	<i>Regression analysis explaining item difficulty in the political knowledge test with principal components of everyday and academic word frequency</i>	145

Anhang Beitrag III

Table B1.	<i>Correlations and descriptive statistics</i>	157
Table B2.	<i>Regression analysis explaining item difficulty in the political knowledge test with different word exclusion criteria</i>	157
Table C1.	<i>Regression analysis explaining item difficulty in the political knowledge test with residualized academic word frequency variables</i>	158
Table C2.	<i>Regression analysis explaining item difficulty in the political knowledge test with residualized everyday word frequency variables</i>	158

Table C3.	<i>Computation of β's and R^2 given the correlation (Friedman & Wall, 2005) between everyday word frequency and item difficulty ($r_{y1} = -.34$), academic word frequency and item difficulty ($r_{y2} = -.09$), and between everyday and academic word frequency ($r_{12} = .77$).....</i>	159
-----------	---	-----

Beitrag V (weiterführende Analysen)

Table 1.	<i>Means, standard deviations, and correlations</i>	208
Table 2.	<i>Pairwise t-tests of average differences across quartiles within and across grade levels</i>	210
Table 2.	<i>Results of the structural equation models</i>	211

Abbildungsverzeichnis

Beitrag I

Abbildung 1.	<i>Vier Bereiche von Civic Literacy</i>	71
Abbildung 2.	<i>Konzeption des Wissenstests zu Civic Literacy</i>	73
Abbildung 3.	<i>Beispiel-Items zum Wissenstest</i>	75
Abbildung 4.	<i>Ergebnisse des Expertenratings</i>	77
Abbildung 5.	<i>Rahmenmodell zu Civic Literacy</i>	81

Beitrag II

Figure 1.	<i>Civic Literacy Framework</i>	99
Figure 2.	<i>Confirmatory factor analysis (adjusted model)</i>	108
Figure 3.	<i>Adjusted Civic Literacy Framework</i>	112

Anhang Beitrag II

Figure B1.	<i>Confirmatory factor analysis (model 1)</i>	124
Figure B2.	<i>Confirmatory factor analysis (model 6)</i>	125

Beitrag III

Figure 1.	<i>Example items in original German (top) and English (bottom)</i>	135
Figure 2.	<i>Bivariate distribution of word frequency in everyday and academic settings</i>	139
Figure 3.	<i>Predicted item difficulty relative to average word frequency in everyday and academic settings</i>	144

Anhang Beitrag III

Figure A1.	<i>Distribution of item difficulty</i>	156
Figure C1.	<i>Graphical display of β's and R^2 given the correlation between everyday word frequency and item difficulty ($r_{y1} = -.34$), academic word frequency and item difficulty ($r_{y2} = -.09$), as well as all possible values for the correlation between everyday and academic word frequency (r_{12})</i>	160

Beitrag IV

Figure 1.	<i>Direct associations between civic education teaching quality facets and willingness to participate</i>	175
-----------	---	-----

Figure 2.	<i>Associations between civic education teaching quality facets and willingness to participate mediated by political interest and political knowledge. Note: Only significant paths are displayed</i>	177
Figure 3.	<i>Associations in 7th and 10th grade between civic education teaching quality facets and willingness to participate mediated by political interest and political knowledge.</i>	179
 Beitrag V (weiterführende Analysen)		
Figure 1.	<i>The Dunning-Kruger effect in Grade 7 and Grade 10</i>	209
 Anhang Beitrag V (weiterführende Analysen)		
Figure A1.	<i>Density distributions of z-standardized responses for objective political knowledge and subjective political knowledge</i>	224
Figure A2.	<i>Bivariate relationship between motivational quality and the difference between subjective and objective political knowledge.....</i>	224
Figure A3.	<i>Bivariate relationships between cognitive activation and objective as well as subjective political knowledge</i>	225

1 Einleitung

Bildungsstandards sind ein vielfach genutztes Instrument, mit dem festgelegt wird, welche Kompetenzen Schüler*innen zu einem bestimmten Zeitpunkt in ihrer Bildungslaufbahn erworben haben sollen (Köller, 2020). Large-Scale Assessments (LSA) können wichtige Erkenntnisse darüber liefern, inwiefern Schüler*innen die formulierten Bildungsstandards erreichen und wie leistungsfähig Bildungssysteme sind (McElvany, 2022). Zwar sind LSAs nicht neu, jedoch haben sie in den vergangenen Jahren weiter an Bedeutung gewonnen (für eine bibliometrische Analyse siehe Hernández-Torrano & Courtney, 2021). Insbesondere international vergleichende LSAs (ILSA) sind ein wichtiges Hilfsmittel, um Erfolg und Effektivität von Bildungssystemen vergleichen zu können (Prenzel, 2018; siehe jedoch auch Johansson, 2016).

Besonders bekannte ILSAs (für eine Analyse von Online-Suchanfragen siehe Jerrim, 2023) sind das von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) durchgeführte *Programme for International Student Assessment (PISA)* sowie die von der *International Association for the Evaluation of Educational Achievement (IEA)* verantworteten *Progress in International Reading Literacy Study (PIRLS)*, *Trends in International Mathematics and Science Study (TIMSS)* und *International Civic and Citizenship Education Study (ICCS)*. Nicht selten dienen die Ergebnisse von ILSAs als Grundlage für bildungspolitische Entscheidungen und Reformen (z. B. PISA: Hopfenbeck et al., 2018; ICCS: Malak-Minkiewicz & Torney-Purta, 2021).

In Deutschland sind nach dem sogenannten *PISA-Schock* Bildungsstandards sowie LSAs in den Fokus der Bildungsadministration und -forschung gerückt (Klieme et al, 2003). Während 2005 für einige Schulfächer (Mathematik, Deutsch, erste Fremdsprache, Biologie, Chemie und Physik) verbindliche nationale Bildungsstandards vorlagen und verabschiedet wurden, oblag es den Fachverbänden und –didaktiken der übrigen Fächer, selbst geeignete Bildungsstandards zu erarbeiten (Köller, 2020). Einer der Fachbereiche für den bis heute keine national verbindlichen Bildungsstandards festgelegt wurden, ist die politische Bildung. Zwar gibt es viele Gründe, warum nationale Bildungsstandards für die politische Bildung problematisch sein könnten (Tampio, 2017), zumal offen ist, inwiefern Bildungsstandards die Leistungen der Schüler*innen und den Unterricht verbessern (Maag Merki, 2016). Allerdings sind Bildungsstandards Grundlage für das nationale Bildungsmonitoring (Kuhn, 2014, McElvany, 2022), sodass befürchtet werden muss, dass für die politische Bildung als Fachbereich ohne Bildungsstandards vergleichsweise wenig empirische Evidenz vorliegt und keine Möglichkeit

zur Reflexion eines „Ist-Soll-Vergleichs“ besteht (siehe Wirkungsmodell nach Maag Merki, 2016, S. 159).

Ein insbesondere im internationalen Kontext gebräuchlicher Begriff, um politische und gesellschaftliche Kompetenz zu beschreiben ist ‚Civic Literacy‘ (z. B. Mirra & Garcia, 202; van Helvoort, 2019). Eine mögliche Definition von Civic Literacy schließt Wissen und Fähigkeiten für informierte Entscheidungen sowie weitere Handlungsdispositionen in Form von Motivationen, Einstellungen und Volitionen ein (Kultusministerkonferenz 2018; Milner 2002; Partnership for 21st Century Skills 2009).

Mit dieser Dissertation soll ein Beitrag zur empirischen Beforschung politischer Bildung und Civic Literacy in der Sekundarstufe I geleistet werden, indem ein Modell entwickelt und erprobt wurde, welches die systematische Erfassung von Civic Literacy ermöglichen soll. Hierfür wird folgenderweise vorgegangen. Im ersten Kapitel wird der theoretische Rahmen der Dissertation erläutert (Kapitel 1). Hierbei wird auf die Bedeutung von Bildung für die Demokratie eingegangen (Kapitel 1.1) sowie die Entwicklung politischer Bildung und relevante Akteur*innen aus nationaler und internationaler Perspektive beschrieben (Kapitel 1.2). Es folgt eine Übersicht zu Herausforderungen politischer Bildung in der Sekundarstufe I (Kapitel 1.3), ein Vergleich nationaler und internationaler Kompetenzmodelle beziehungsweise *Assessment Frameworks* (Kapitel 1.4) und eine Zusammenfassung testtheoretischer Herausforderungen (Kapitel 1.5). Zum Abschluss des ersten Kapitels werden die übergeordneten Forschungsfragen beschrieben (1.6) und die Beiträge der kumulativen Dissertation zusammengefasst (1.7). Das zweite Kapitel enthält die Einzelbeiträge dieser Dissertation und beginnt mit einem Beitrag über die Konzeptualisierung und Operationalisierung eines Civic Literacy Kompetenzmodells für die politische Bildung in der Sekundarstufe I (Kapitel 2.1). Hiernach wird das entwickelte Kompetenzmodell anhand unterschiedlicher Gütekriterien und auf Basis empirischer Daten überprüft (Kapitel 2.2). Im Folgenden wird der im Kompetenzmodell enthaltene Wissenstest daraufhin untersucht, inwiefern Worthäufigkeiten die Schwierigkeit von einzelnen Aufgaben beeinflussen (Kapitel 2.3). Das Kapitel schließt mit einem Beitrag zum Zusammenhang zwischen unterschiedlichen Unterrichtsmerkmalen und verschiedenen Bereichen des Kompetenzmodells (Kapitel 2.4). Das dritte Kapitel enthält weiterführende Analysen zum Zusammenhang zwischen subjektivem und objektivem politischen Wissen (Kapitel 3). Im vierten Kapitel folgt eine generelle Diskussion der Ergebnisse, im Zuge derer die unterschiedlichen Beiträge in Verbindung zueinander gesetzt und deren Ergebnisse diskutiert (Kapitel 4.1), die

Stärken und Limitation der Arbeit dargelegt (Kapitel 4.2) und etwaige Implikationen für Forschung und Praxis behandelt werden (Kapitel 4.3). Abschließend wird ein Fazit gezogen (Kapitel 4.4).

1.1 Demokratie und Bildung

Die Demokratie gilt als wichtiges Instrument der kollektiven Entscheidungsfindung und dient vielfach als Ideal für Herrschafts- und Gesellschaftssysteme (Edelstein & Fauser, 2001). Demokratien können jedoch entscheidungsträge sein und sind mit vergleichsweise hohem Aufwand für die Bürger*innen verbunden (Brenner, 2016). Zudem sind die Erreichung und das erfolgreiche Fortbestehen demokratischer Systeme keine Selbstverständlichkeit. Tatsächlich sind viele existierende als ‚demokratisch‘ geltende Systeme nur sogenannte Polyarchien, das heißt eine Approximation des Idealtyps einer Demokratie (Dahl, 1989). Nichtsdestotrotz gilt die Demokratie als das zumindest theoretisch überlegene System, wenn es darum geht, grundlegende Freiheiten und Rechte zu sichern (Dahl, 1989).

Weltweit stehen Demokratien und mit ihnen das demokratische Ideal der Gesellschaft und der Politik jedoch unter Druck. In 2022 verzeichnete der Bertelsmann Transformationsindex (BTI) erstmals seit 2004 mehr autokratische als demokratische Staaten (Hartmann & Thiery, 2022). Von den 137 analysierten Staaten wurden 67 als demokratisch und 70 als autokratisch regiert eingestuft. Insgesamt neun Demokratien haben im vergangenen Jahrzehnt so deutliche Rückschritte gemacht, dass sie laut Bertelsmann Stiftung mittlerweile als gemäßigte Autokratien eingestuft werden müssen (Hartmann & Thiery, 2022). Seit 2010 sinkt das globale Demokratieniveau und erreichte 2021 den niedrigsten Wert seit 1989 (Boese et al, 2022). Von der negativen Entwicklung sind Staaten in allen Regionen der Welt betroffen.

Auch das politische System der BRD steht vor der Herausforderung, dass die Teilnehmeraten in verschiedenen (konventionellen) Beteiligungsformaten in den vergangenen beiden Jahrzehnten zurückgegangen sind (z. B. Glavanovits et al., 2019). Beispielsweise haben in Deutschland vier der sechs mitgliederstärksten Parteien seit 1990 mehr als die Hälfte ihrer Mitglieder verloren (Winter, 2021). Ein ähnlicher Trend zeigt sich für die Wahlbeteiligung bei Bundestagswahlen. Die Höchstwerte aus den 1970er mit über 90 Prozent Wahlbeteiligung werden mittlerweile deutlich unterschritten (Bundeswahlleiterin, 2023). Nachdem die Wahlbeteiligung 2009 auf den historischen Tiefstwert von 70,8 Prozent sank, konnte zuletzt jedoch ein leichter Aufwärtstrend festgestellt werden und die Wahlbeteiligung stieg bei den Wahlen 2017 und 2021 wieder auf ca. 76 Prozent (Bundeswahlleiterin, 2023). Die sinkende

Wahlbeteiligung ist jedoch kein allein deutsches Problem. Weltweit stehen gefestigte Demokratien vor der Herausforderung sinkender Wahlbeteiligung (Kostelka, 2017).

Einer der Gründe für die niedrige Wahlbeteiligung ist, dass insbesondere junge Menschen seltener Gebrauch von ihrem Wahlrecht machen. Ein Kohortenvergleich der Daten der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) von 1984 bis 2018 ergab, dass Menschen im Alter von 18 bis 34 Jahren im Verlauf der Zeit seltener angaben, zur Wahl zu gehen (Holderberg, 2021). Auch die Analyse von Wahlen in 22 Europäischen Ländern zwischen 1999 und 2002 zeigte, dass die Gruppe der wahlberechtigten Menschen unter 25 eine durchschnittlich 19 Prozent niedrigere Wahlbeteiligung hatte als die Gesamtbevölkerung (Fieldhouse et al., 2007). Umso wichtiger ist es, junge Menschen zur Mündigkeit zu erziehen und ihnen die für demokratisches Handeln und politisches Partizipieren nötigen Kompetenzen (d. h. Civic Literacy) zu vermitteln (Hahn-Laudenberg, 2022).

Schon lange ist klar, dass Bildung ein wichtiges Instrument ist, um demokratische Strukturen zu stärken und zu erhalten (Dewey, 1916; Haarmann et al., 2020). Auch empirisch konnte mithilfe eines Ländervergleichs belegt werden, dass Demokratien von den Einstellungen und dem Engagement ihrer Bürger*innen abhängen und demokratische Institutionen alleine nicht genügen (Alma & Verba, 1963). Konkret ist die Unterstützung der Werte, Strukturen und Normen sowie der relevanten Gesetzgebung eines politischen Systems gemeint. Bezogen auf die BRD schließt dies die Unterstützung der freiheitlich demokratischen Ordnung, des Grundgesetzes und der Menschenrechte ein (Gabriel, 2001).

Gut ausgebildete Bürger*innen sind ein wichtiger Faktor für die Erreichung und Erhaltung einer Demokratie. Eine internationale Vergleichsstudie mit Paneldaten aus 95 autoritären Regierungszeiten zwischen 1970 und 2008 konnte zeigen, dass ein höheres allgemeines Bildungsniveau auch unter Berücksichtigung der wirtschaftlichen Entwicklung mit Demokratisierungstendenzen zusammenhing (Sanborn & Thyne, 2014). Auf Individualebene wurde zudem gefunden, dass das allgemeine Bildungsniveau und die Wahrscheinlichkeit einer Teilnahme an politischen Wahlen positiv zusammenhängen (Sondheimer & Green, 2010). Anhand von drei Quasi-Experimenten, in denen die Folgen von Bildungsreformen untersucht wurden, fanden Sondheimer & Green (2010), dass die Reformen zu höheren Schulabschlussquoten und auch zu höheren Wahlbeteiligungen führten. Paneldaten aus England legen zudem nahe, dass die Aufnahme eines universitären Studiums die Wahrscheinlichkeit, an einer politischen Wahl teilzunehmen, erhöht (Yang & Hoskins, 2020).

Es existiert jedoch auch die Annahme, Bildung sei nicht kausal für politische Teilhabe verantwortlich (*cause*-Perspektive: z. B. Yang & Hoskins, 2020), sondern nur ein *proxy* und

individuelle Attribute wie Intelligenz und Persönlichkeit würden sowohl Bildungsentscheidungen als auch Partizipationsentscheidungen beeinflussen (*education-as-a-proxy*-Perspektive: z. B. Burden et al., 2020). Eine weitere Möglichkeit könnte sein, dass unterschiedliche Bildungswege mit unterschiedlichen sozialen Netzwerken einhergehen, die wiederum Partizipationstendenzen beeinflussen (*relative-education*-Perspektive: z. B. Aars & Christensen, 2020). Ein Systematic Literature Review der *cause*, *education-as-a-proxy* und *relative education* Perspektiven legt jedoch nahe, dass insbesondere die *cause*-Perspektive den Zusammenhang zwischen Bildung und politischer Partizipation erklärt (Le & Nguyen, 2021; Campbell, 2019).

Durch Bildung werden Fähigkeiten erlernt, die notwendig sind, um politische Entscheidungen zu treffen und die Auswirkungen von politischen Maßnahmen abzuschätzen (Glaeser et al., 2007). Eine gut ausgebildete Bevölkerung ist auch besser in der Lage, ihre Interessen zu artikulieren und zu verteidigen und sich gegen Missbrauch und Ungerechtigkeit zu wehren. Anhand der Analysen von 51 Sachabstimmungen in der Schweiz wurde beispielsweise gezeigt, dass Menschen mit einem höheren Bildungsstand öfter politisch partizipieren und ihre Abstimmungsentscheidungen besser mit ihren persönlichen Präferenzen übereinstimmen (Hedinger, 2018). Umgekehrt können Bildungsdefizite dazu führen, dass Menschen leichter manipulierbar und empfänglicher für populistische Propaganda sind (z. B. Tsatsanis et al., 2018).

1.2 Politische Bildung in Deutschland und im internationalen Kontext

Politische Bildung ist ein wichtiger Bestandteil offener und liberaler Gesellschaften, da sie junge Menschen auf ein Leben in der Demokratie vorbereitet und so die Demokratie nachhaltig stärkt. In Deutschland (Kultusministerkonferenz, 2018) wie auch in anderen Ländern (z. B. Gould et al., 2011) hat politische Bildung das Ziel, politische Dispositionen wie politisches Wissen, politisches Interesse und demokratische Einstellungen zu vermitteln. Lernangebote im schulischen und außerschulischen Bereich werden durch Bundesprogramme zur Demokratiestärkung flankiert (z. B. „Demokratie leben!“, „Land.schafft.Demokratie“).

In Deutschland gilt, dass politische Bildung beziehungsweise Demokratiebildung (für einen Vergleich siehe Fischer & Reinhardt, 2022) nicht ausschließlich affirmativ, sondern auch kritisch sein muss, damit junge Menschen lernen, „die gegebene Ordnung zu verstehen, zu reflektieren, zu verändern, zu kritisieren und so zu gestalten, wie sie es selbst für adäquat halten“ (Kenner & Lange, 2020, S. 49). Die Notwendigkeit zur kritischen Mündigkeit (auch gegenüber der Demokratie selbst) und der von der Kultusministerkonferenz an die Schulen formulierte

Auftrag (Kultusministerkonferenz, 2018), ein Wertesystem zu vermitteln, welches freiheitlichen und demokratischen Grund- und Menschenrechten entspricht, schafft jedoch ein Spannungsverhältnis für die Schulpraxis (Gerdes, 2022). Auch wenn ein Auflösen dieses Spannungsverhältnis nicht unmittelbar möglich ist, sollten Bildungsakteur*innen hierfür sensibilisiert werden (May, 2021).

Ein barrierefreier und fairer Zugang zu politischen Informationen und Beteiligungsformaten sind ein wichtiger Bestandteil nachhaltiger Demokratieförderung (Aktionsrat Bildung, 2020). Hieraus ergibt sich eine besondere Bedeutung der Schule, da über sie alle zukünftigen Bürger*innen einer Gesellschaft systematisch erreicht werden (Abs & Hahn-Laudenberg, 2017). Als Mikromodell der Gesellschaft hat die Schule die Aufgabe, Räume zu schaffen, die jungen Menschen die Möglichkeit geben, demokratische Verhaltensweisen einzuüben und politische Meinungen zu diskutieren (Aktionsrat Bildung, 2020). Die Schule sollte dabei kein Ort der prinzipiellen Wertfreiheit sein, an dem politische Bildung als neutrale und institutionenorientierte Staatskunde gelehrt wird (Grob, 2009). Stattdessen sollte die Schule ein pädagogisches Handeln kultivieren, das sich aus den Grund- und Menschenrechten des Grundgesetzes ableiten lässt (Kultusministerkonferenz, 2018).

Tagtäglich gelingt an deutschen Schulen die Vermittlung wichtiger politischer und demokratischer Fähigkeiten im Fachunterricht sowie in fächerübergreifenden Lernsettings. Allerdings gibt es auch an deutschen Schulen Optimierungsbedarfe. Ergebnisse der international vergleichenden ICCS 2016 zeigten, dass das mittlere politische Wissen von Schüler*innen in Nordrhein-Westfalen (NRW)¹ statistisch signifikant unterhalb der europäischen Vergleichsgruppe und Ländern wie zum Beispiel Kroatien, Slowenien und der Russischen Föderation lag (Abs & Hahn-Laudenberg, 2017). Zudem stand die Testleistung der Schüler*innen in NRW stärker als in anderen Ländern mit herkunftsbedingten Merkmalen (Migrationshintergrund und kulturelles Kapital) in Zusammenhang.

International werden neben *civic literacy* und *citizen competence* vor allem die Begriffe *critical thinking* und *critical inquiry* genutzt. Unter anderem stehen hier der Unterschied zwischen kritischen Unterrichtsinhalten und kritischem Denken sowie Unterrichtsmethoden zum Erlernen von *critical thinking* im Fokus (Bermudez, 2015; Yang & Chung, 2009). Im angelsächsischen Raum hat die politische Bildung zudem häufig einen handlungsvorbereitenden

¹In Deutschland nahm ausschließlich Nordrhein-Westfalen teil. Die Teilnahmequoten auf Schul- und/oder Schülerebene in Nordrhein-Westfalen erreichten nicht die internationalen Vorgaben, wodurch die Repräsentativität nicht abgesichert ist. Aufgrund des international vergleichenden Charakters wird im Test kein Bezug zum deutschen politischen System und dessen spezifischen institutionellen Regeln und Strukturen, Konfliktlinien, Aushandlungsprozessen und Politikfeldern hergestellt.

Charakter und enthält Elemente der *service-learning* Methode (Jerome, 2012) und der *action civics* (Lau, 2017). Eine Gemeinsamkeit ist jedoch, dass die politische Bildung tief in den politischen und historischen Kontext eines Landes eingebettet ist (Torney-Purta et al., 1999).

1.2.1 Ausgangspunkt und Entwicklung

Ursprünglich war politische Bildung vor allem ein Instrument, um Menschen zu loyalen und gleichgesinnten Bürger*innen zu erziehen, die das bestehende politische System akzeptieren ohne es zu hinterfragen (Bromley & Mäkinen, 2011). Erst im Laufe des 20. Jahrhunderts wurden ein positiveres Bild von Diversität und Aspekte der Mündigkeit beziehungsweise *critical thinking* in die schulische politische Bildung integriert (Crittenden & Levine, 2018). Auch Multikulturalität ist ein eher junges Thema in der politischen Bildung und unterschiedlich stark in verschiedenen Ländern ausgeprägt (Bromley & Mäkinen, 2011).

Die historische Bedeutung der politischen Bildung in Deutschland reicht bis in das frühe 20. Jahrhundert zurück und ist eng mit den politischen, gesellschaftlichen und bildungspolitischen Entwicklungen Deutschlands verbunden. Zu Beginn des 20. Jahrhunderts wurden erste Versuche unternommen, die schulische politische Bildung zu einem eigenständigen Unterrichtsfach im Sinne einer ‚Staatsbürgerkunde‘ weiterzuentwickeln, welches die junge Generation den Glauben an Kaiser und Gott lehren und vor den Gefahren der Sozialdemokratie warnen sollte (Hafeneger, 2018). Im Kontext der Weimarer Republik wurde das schulische Angebot vor allem durch außerschulische Angebote mit einem breiten Trägerspektrum erweitert. Hierin enthalten waren unter anderem selbstorganisierte Lernformen in der Reformpädagogik und Arbeiterbewegung, subjektorientierte Reformansätze sowie autoritär und national-politische Erziehungsmethoden (Hafeneger, 2018). Zwar fand die nationalsozialistische Erziehungsdoktrin einzelne Anknüpfungspunkte in den existierenden Konzeptionen politischer Bildung, dennoch stellte die politische Bildung des Nationalsozialismus einen deutlichen Bruch mit den bis zu dem Zeitpunkt bedeutenden Erziehungsmethoden der abendländischen Welt dar. Im Mittelpunkt der Erziehung stand der Körper. Charakter und Willensstärke waren zweitrangig, während Geist und Intellekt nur von nachgeordneter Bedeutung waren (Hafeneger, 2018). Ein zweites zentrales Element der nationalsozialistischen Erziehung war die Rassenlehre (Kuhn et al., 1993).

Nach dem Zweiten Weltkrieg und dem Ende des Nationalsozialismus wurde die politische Bildung als Reaktion auf die Erfahrungen der Diktatur und des Totalitarismus als wichtiger Bestandteil des demokratischen Neuaufbaus betrachtet und vor allem durch die Re-education Politik der Alliierten bestimmt (Sander, 2021a). Nach Gründung der BRD und der Deutschen

Demokratischen Republik (DDR) entwickelte sich die politische Bildung jedoch in unterschiedliche Richtungen.

In der DDR wurde das Fach Staatsbürgerkunde eingeführt. Quantitativ fiel das Fach vergleichsweise kaum ins Gewicht, jedoch wurde das Fach bis zuletzt als Schlüsselfach und unersetzliches Instrument für die sozialistische Erziehung und die Vermittlung der Staatsideologie angesehen (Grammes et al., 2006). Ähnlich zum Erziehungsideal des Nationalsozialismus galt die politische Erziehung in der DDR vordergründig als Instrument der Herrschaftsstabilisierung, mit der Bindungen an das politische System und dessen Ideologie gesichert werden sollten (Hafeneger, 2018). Der Erziehungsanspruch bezog sich in der DDR nicht nur auf die schulische Bildung, sondern durchzog die gesamte Gesellschaft. Die politische Erziehung galt als Lebensprinzip, mit dem Ziel der Entwicklung einer sozialistischen Gesellschaft und Persönlichkeit (Grammes et al., 2006). Inhaltlich war insbesondere der Marxismus-Leninismus von Bedeutung, der auch als Handlungsgrundlage für die Politik der DDR galt (Grammes et al., 2006).

In der BRD unterstützten die Alliierten die Entwicklung demokratischer Bildungsinstitutionen, die curriculare Verankerung politischer Bildung in Schulen und die Förderung außerschulischer Angebote. Die politische Bildung wurde in den 1950er und 1960er Jahren vor allem durch die Gründung von Institutionen wie der Bundeszentrale für politische Bildung (bis 1963 Bundeszentrale für Heimatdienst) und den Landeszentralen für politische Bildung vorangetrieben (Golz & Kost, 2021). Diese Institutionen sollten politisches Wissen vermitteln, politische Diskussionen fördern und zur demokratischen Bewusstseinsbildung beitragen (Kuhn et al., 1993). Die Auffassung, politische Bildung sei ein Instrument zur Legitimierung der Herrschaft und der bestehenden Weltanschauung, wurde durch eine auf Mündigkeit und Selbstbestimmung zielende demokratische politische Bildung abgelöst (Hafeneger, 2018). Während der 1960er und 1970er Jahre gewann die politische Bildung an Bedeutung, wurde im Rahmen der „Didaktischen Wende“ professionalisiert sowie stärker wissenschaftlich ausgerichtet und arbeitete zunehmend zielgruppenorientiert (Sander, 2021). Thematisch wurden der politische Diskurs und die politische Bildung durch die Folgen der 68er-Bewegung, den kalten Krieg und die Auseinandersetzung mit der nationalsozialistischen Vergangenheit geprägt (Hafeneger, 2018). Hinzu kamen die Suche nach und der Fokus auf neue Formen politischer Partizipation, in deren Mittelpunkt das Erstarken zivilgesellschaftlicher (Protest-)Bewegungen stand (Hafeneger, 2018). Insbesondere der Ausbau der Kernenergie, der NATO-Doppelbeschluss sowie Themen der Geschlechtergerechtigkeit wurden vielfach von breit geführten gesellschaftlichen Debatten und Protesten begleitet (Gagel, 2005). Eine bis heute in der

politischen Bildung bedeutsame Rolle spielt der 1976 von Vertreter*innen der politischen Bildungspraxis, -forschung und -administration vereinbarte *Beutelsbacher Konsens*, der die drei Kernprinzipien *Überwältigungsgebot*, *Kontroversitätsgebot* und *Interessenorientierung* enthält (z. B. Frech & Richter, 2017).

Mit dem Mauerfall und der Wiedervereinigung Deutschlands im Jahr 1990 eröffneten sich neue Herausforderungen und Chancen für die politische Bildung. Faktisch wurden in den 1990er Jahren westdeutsche Konzepte und Denkansätze der politischen Bildung für das wiedervereinigte Deutschland übernommen (Sander, 2021a). Die Auseinandersetzung mit der deutschen Teilung und der deutschen Geschichte wurden wichtige Themen in der politischen Bildungsarbeit (Sander, 2021a). Zudem wurde inhaltlich zunehmend ein Diversity-Ansatz verfolgt, der Aspekte wie interkulturelle Bildung, Geschlechterperspektiven und globales Lernen integrierte (Achour, 2021). Zu Beginn der 2000er Jahre wurde das bildungspolitische Geschehen vor allem durch das enttäuschende Abschneiden deutscher Schüler*innen in der PISA 2000 Studie bestimmt. Das in der Folge entstandene Bestreben zur Einführung von domänenspezifischen Bildungsstandards sowie die damit einhergehende *Empirische Wende* waren und sind insbesondere für die *schulische* politische Bildung wichtige Impulsgeber (Weißeno, 2018). In den letzten Jahren erfuhr vor allem die politische Jugendbildung aufgrund von sozialen Bewegung wie *Fridays for Future*, *Black Lives Matter* und *Pulse of Europe* eine erhöhte Aufmerksamkeit (Kenner, 2021).

In den USA wurde ab dem Beginn des 20. Jahrhunderts vermehrt über die Bedeutung von politischer Bildung (*civic education*) für die Demokratie gesprochen (Dewey, 1916). Es dauerte jedoch bis in die 1960er Jahre bis Forschende (insbesondere aus den USA und Westeuropa) sich systematisch und empirisch mit der Frage auseinandersetzten, wie junge Menschen politische und demokratische Kompetenzen erwerben (Hahn, 2010). Nachdem in einer vielfach beachteten Studie jedoch vermeintlich keine Zusammenhänge zwischen der Teilnahme am Politikunterricht und politischem Wissen gefunden wurden (Langton & Jennings, 1968), flachte das Interesse an politischer Bildung insbesondere in den USA, aber auch international, wieder ab (Campbell, 2019).

In ihrer Studie haben Langton und Jennings die Daten von 1.669 *high school* Schüler*innen aus 97 US-amerikanischen Schulen genutzt und haben insgesamt nur einen kleinen und statistisch nicht signifikanten Zusammenhang für die Teilnahme an *civic education* Kursen in den Jahrgangsstufen 10, 11 und 12 und dem politischen Wissen zur Zeit der Erhebung festgestellt. Für die Sub-Gruppe schwarzer Schüler*innen, die durchschnittlich schlechter abschnitten als ihre weißen Mitschüler*innen, stellten Langton und Jennings jedoch sehr wohl

einen signifikanten und positiven Zusammenhang fest. Weniger kognitive Maße wie das politische Interesse und politische Einstellungen wurden in der Studie nicht betrachtet. Nichtsdestotrotz bestand drei Jahrzehnte lang die weit verbreitete Meinung, dass Politikunterricht nur geringe beziehungsweise keine Auswirkungen auf das politische Wissen und die politischen Orientierungen von Schüler*innen hat (Galston, 2007). Eine einige Jahre später durchgeführte Studie mit Daten des National Assessment of Educational Progress (NAEP), in der US-amerikanische Schüler*innen der 12. Klassenstufe befragt wurden, zeigte jedoch, dass sowohl Umfang als auch Aktualität (*amount* und *recency*) des Politikunterrichts mit politischem Wissen sowie dem Vertrauen in politische Institutionen und das politische Effektivitätsgefühl zusammenhängen (Niemi & Junn, 1998). Die Studie von Niemi und Junn sowie die politischen Entwicklungen infolge der Erosion der bipolaren Weltordnung sorgten für ein erneutes internationales Forschungsinteresse zu politischer Bildung (Torney-Purta et al., 1999a). Seit dem Ende der 1990er Jahre sind vor allem die Civic-Studien der IEA eine wichtige Ressource für die Beforschung von politischer Bildung. Im Jahr 1999 führte die IEA die CIVED-Studie durch, die bis dahin die größte internationale Studie auf diesem Gebiet war. 2009 folgte die ICCS, die auch 2016 und 2022 durchgeführt wurde und bei der Schüler*innen im Alter von ca. 13 Jahren zu ihrer politischen Bildung befragt wurden (Schulz et al., 2016).

1.2.2 Relevante Akteur*innen

Insbesondere in demokratischen Gesellschaften gibt es eine Vielzahl von Institutionen, Verbänden, Organisationen und Trägern, die politische Bildungsangebote anbieten (Golz & Kost, 2021). National wie auch international kann zwischen außerschulischen und schulischen Akteur*innen der politischen Bildung unterschieden werden. Beide arbeiten auf der Grundlage politischer Vorgaben und rechtlicher Rahmenbedingungen, unterscheiden sich jedoch mit Blick auf die Strukturen und Lernformen (Hafeneger, 2021).

In der BRD wird die schulische politische Bildung in den Landesverfassungen, den Schulgesetzen der Länder, den Richtlinien der Kultusminister*innen sowie den Stundentafeln für die unterschiedlichen Schulformen und -stufen geregelt (Hafeneger, 2018). Politische Bildung ist ein fester Bestandteil des Schulunterrichts in Deutschland. Nicht nur in den originären Bezugsfächern wie dem Politikunterricht oder dem Sozialkundeunterricht werden politische Themen verhandelt, sondern auch in den weiteren Fächern des sozialwissenschaftlichen Spektrums (z. B. Geschichte und Erdkunde) sowie in selteneren Fällen auch in anderen Fächern (Kultusministerkonferenz, 2018). Wichtige Akteurinnen sind daher die Bildungsadministratio-
nen der Bundesländer sowie die Kultusministerkonferenz. Darüber hinaus sind auch die

Schulen selbst wichtige Akteurinnen, da das Schulgesetz ihnen in Form von Kontingent- und Bereichsstunden gewisse Freiheiten bei der Unterrichtsplanung einräumt (Reichhart, 2018). Schließlich sind Lehrkräfte zentrale Akteur*innen, da sie zumindest theoretisch (Schneider & Gerold, 2018) erheblichen Einfluss auf die Ausgestaltung des Unterrichts und somit auf die Qualität politischer Lerngelegenheiten nehmen können (Reichhart, 2018).

Die außerschulische politische Bildung (hiermit sind auch Ganztagsangebote von Schulen gemeint) wird in der BRD in den Landesgesetzen geregelt und durch Vorgaben im Kinder- und Jugendgesetz gerahmt (Hafeneger, 2021). Sie gehört im engeren Sinne zu den Schwerpunkten der Jugendarbeit, soll an die „Interessen junger Menschen anknüpfen und von ihnen mitbestimmt und mitgestaltet werden“ und hat zum Ziel „zur Selbstbestimmung zu befähigen und zu gesellschaftlicher Mitverantwortung und zu sozialem Engagement anzuregen und hinführen“ (§ 11 Absatz 1 Satz 2 SGB VIII). Relevante Akteur*innen umfassen zivilgesellschaftlich, konfessionell und gewerkschaftlich geprägte Bildungsträger*innen sowie kommunale Bildungsstätten, die Bundeszentrale und die Landeszentralen für politische Bildung, Volkshochschulen, Jugendverbände und Stiftungen, Bildungswerke und zahlreiche Initiativen (Hafeneger, 2021). Insbesondere die Bundeszentrale sowie die Landeszentralen politischer Bildung sind wichtige außerschulische Institutionen der politischen Bildung in staatlicher Trägerschaft und sollen politische Themen möglichst objektiv und offen für kontroverse Meinungen behandeln (Golz & Kost, 2021). Ausgestattet mit der Aufgabe, politisches Wissen zu vermitteln, politische Diskussionen zu fördern und zur demokratischen Bewusstseinsbildung beizutragen, bieten sie eine Vielzahl von Publikationen, Veranstaltungen und digitalen Angeboten an (Golz & Kost, 2021).

Im internationalen Kontext wird häufig von *civic education* gesprochen, wenn es um politische Bildung geht. Der Begriff wurde vor allem in den angelsächsischen Ländern geprägt und bezieht sich zumeist auf die *schulische* politische Bildung (Frank, 2005). In den USA werden die Lehrpläne ebenfalls durch die jeweiligen Bundesstaaten und darüber hinaus durch die Schuldistrikte rechtlich koordiniert (Hahn, 2008). Neben dem Fachunterricht spielen auch im Kontext der Schule organisierte außerunterrichtliche Angebote (z. B. Sportmannschaften oder Kunst-AGs) eine Rolle (Darling et al., 2017). Ein weiteres wichtiges Element an der Schnittstelle zwischen außer- und innerschulischer politischer Bildung in den USA ist die *service-learning* Methode, bei der Schüler*innen eine gewisse Anzahl von Stunden pro Jahr Freiwilligenarbeit ableisten (Jerome, 2012).

Außerdem bedeutsam sind transnationale Bildungsakteur*innen, die im außerschulischen Kontext agieren. Beispiele hierfür sind die OECD, die Weltbank oder Stiftungen wie die

Bertelsmann Stiftung und die Bill und Melinda Gates Stiftung (Gogolin, 2020). Auch wenn politische Bildung kein originäres Anliegen dieser Akteur*innen ist, spielt sie häufig eine Rolle. Kernthemen wie zum Beispiel *global citizenship* haben oft einen konkreten Bezug zur politischen Bildung (Scheunpflug, 2021). Eine zunehmend wichtigere Rolle spielen zudem transnational agierende Onlineplattformen, die indirekt und ohne dies dem Nutzenden kenntlich zu machen, politische Inhalte ideologisch motiviert und einseitig vermitteln (Waldis, 2020).

Aus der Forschungsperspektive treten unterschiedliche wissenschaftliche Bezugsdisziplinen als wichtige Akteurinnen in Erscheinung. Traditionell bedeutsam sind international wie auch national die Politikwissenschaften, Erziehungswissenschaften und in geringerem Umfang auch die Soziologie und Wirtschaftswissenschaften (Campbell, 2019; Weißeno, 2018). Im Zuge der stärkeren Outputorientierung im Bildungskontext seit den 2000er Jahren und dem damit verbundenen Fokus auf den Kompetenzbegriff hat jedoch auch die Pädagogische Psychologie an Bedeutung gewonnen. In Deutschland hat sich die Politikdidaktik als Leitdisziplin bewährt, die vermehrt als eigene Wissenschaftsdisziplin beschrieben wird (Weißeno, 2018).

1.3 Politische Bildung in der Sekundarstufe I

Die Förderung von politischer Teilhabe erfolgt in der Sekundarstufe I vor allem indirekt über die Vermittlung von für die Partizipation nötige Kompetenzen (Oberle & Schwanholz, 2023). Eine Studie mit $N = 2.931$ Belgischen Schüler*innen der 8. Klassenstufe konnte beispielsweise zeigen, dass der Zusammenhang zwischen Unterricht und politischer Partizipation größtenteils über das politische Interesse und das politische Effektivitätsgefühl mediiert wird (Maurissen, 2020).

In der BRD variieren die curricularen Rahmenbedingungen und konkreten Zielsetzungen für die politische Bildung der Sekundarstufe I je nach Bundesland und Schulform. Insbesondere in Bezug auf Benennung und Umfang bestehen Unterschiede. Dennoch gibt es einige allgemeine Prinzipien und Ziele, die in den unterschiedlichen Lehrplänen verankert sind. Mit wenigen Ausnahmen werden Schüler*innen in allen Bundesländern und Schulformen erstmals im Laufe der Sekundarstufe I in einem Fach unterrichtet, in dem es originär um politische Bildung geht (Gökbudak et al., 2021). Als Gelingungskriterien schulischer politischer Bildung sind insbesondere die Quantität von Lerngelegenheiten und der damit verbundene zeitliche Umfang politischer Bildung in der Schule sowie die Qualität von Lerngelegenheiten einschließlich der Rolle der Lehrkraft von Bedeutung (Achour & Wagner, 2019).

1.3.1 *Bezeichnung und Umfang*

Unterschiede zwischen den Bundesländern zeigen sich bereits in der Heterogenität der Bezeichnung des Leitfachs der politischen Bildung. Beispiele hierfür sind die Fächer „Politik“, „Politische Bildung“, „Gemeinschaftskunde“, „Sozialkunde“ oder auch „Gesellschaftslehre“ (Gökbudak et al., 2021). Bei Betrachtung dieser Fächer wird zudem deutlich, dass politische Bildung häufig gemeinsam mit anderen Fächern des sozialwissenschaftlichen Spektrums unterrichtet wird (Ziegler, 2018).

Ein Versuch der Quantifizierung des laut Lehrplänen vorgesehenen zeitlichen Umfangs politischer Bildung in den Bundesländern erfolgt durch das regelmäßig erscheinende *Ranking Politische Bildung* (Gökbudak et al., 2021). Das Ranking soll Transparenz über den aktuellen Stand der schulischen politischen Bildung in Deutschland schaffen. Es werden Indikatoren wie die Stundentafelquote, die Platzierung und die Kontinuität des Politikunterrichts erfasst und ein Vergleich zwischen den Bundesländern und Schulformen vorgenommen (Gökbudak et al., 2021). Die Ergebnisse des Rankings für das Jahr 2021 zeigten, dass Bayern, Thüringen und Rheinland-Pfalz weiterhin die Bundesländer sind, in denen sowohl an Gymnasien als auch an nichtgymnasialen Schulformen am wenigsten politische Bildung vorgesehen ist (Gökbudak et al., 2021). Vergleichsweise viel politische Bildung findet in NRW und Schleswig-Holstein statt, die sowohl im Ranking der Gymnasien sowie der nichtgymnasialen Schulformen jeweils einen der vorderen drei Plätze belegten. Insgesamt muss festgestellt werden, dass sich die curriculare Verankerung stark zwischen den 16 Bundesländern unterscheidet. An bayerischen Gymnasien ist die Stundentafelquote für das Leitfach der politischen Bildung etwa acht Mal niedriger als die Stundentafelquoten in NRW, Hessen oder Schleswig-Holstein (Gökbudak et al., 2021).

Eine inhaltliche Curriculumanalyse für die Sekundarstufe I und II in den sechs bevölkerungsreichsten Bundesländern zeigte zudem, dass die Themen *Politisches System Deutschlands* sowie *Internationale Beziehungen* und *Europäische Union* besonders häufig im Unterricht behandelt werden sollen (Hippe et al., 2020). Die Themen *Extremismus und Multikulturelle Gesellschaft/Migration*, *Ökologische Nachhaltigkeit* und *Geschlechter(politik)* finden hingegen kaum oder gar keine Beachtung in den Curricula. Die Analyse der zu vermittelnden Kompetenzen zeigte zudem, dass es in den curricularen Vorgaben der sechs Bundesländer vor allem um Systemorientierung und weniger um Problemorientierung geht und die fokussierten Themen einer „modernisierten Institutionenkunde“ ähneln (Hippe et al., 2020, S. 7). Zwar sollen in vielen Ländern auch konventionelle Partizipationsmöglichkeiten im Sinne

einer politischen Handlungskompetenz behandelt werden, jedoch liegt auch hier der Fokus auf institutionalisierten Verfahren (Hippe et al., 2020).

1.3.2 Unterrichtsqualität

Neben der Quantität spielt auch die Qualität politischer Lerngelegenheiten eine wichtige Rolle für die Zielerreichung in der schulischen politischen Bildung. Eine Schlüsselrolle nehmen in diesem Zusammenhang die Fachlehrkräfte ein, die einen Einfluss auf die Qualität politischer Lerngelegenheiten nehmen können (Ziegler, 2018). Systematische Studien zur Bedeutung der Unterrichtsqualität bezogen sich bislang jedoch vorrangig auf die Kernfächer und die Fächer des naturwissenschaftlichen Spektrums (Weißeno & Grobshäuser, 2021).

Eine allgemeine, pädagogisch-psychologische Systematisierung der Unterrichtsqualität geht von den drei Basisdimensionen *Klassenführung*, *Kognitive Aktivierung* und *Konstruktive Unterstützung* aus (Klieme, 2022). Die Basisdimensionen können jedoch nicht unmittelbar auf alle Fächer angewendet werden, sondern bedürfen verschiedener Ergänzungen (Praetorius et al., 2020). Mit Ausnahme der Mathematikdidaktik (Kunter & Ewald 2016) wurden die Diskurse zu einer fachspezifischen Unterrichtsqualität in den Fachdidaktiken lange nur cursorisch geführt und Annahmen hierzu selten empirisch überprüft (Praetorius & Gräsel, 2021a).

Trotz der sogenannten *Fachdidaktischen Wende* (Reusser & Pauli, 2021), die die Bedeutung fachspezifischer Aspekte für die Unterrichtsqualität unterstreicht, ergibt sich für grundsätzliche Überlegungen und empirische Befunde innerhalb der Politikdidaktik eine Angebotslücke. So finden sich in einem in 2021 erschienenen Themenheft zur fachspezifischen Unterrichtsqualität Beiträge für die Fächer Deutsch, Englisch, Musik, Kunst, Sport, Religion, Geschichte, Biologie und Mathematik, nicht jedoch für das Fach Politik (Praetorius & Gräsel, 2021b). Nichtsdestotrotz gibt es auch für den Bereich der schulischen politischen Bildung einige empirische Untersuchungen, die sich mit der Bedeutung von einzelnen Unterrichtsqualitätsmerkmalen für den Lernprozess und Lernerfolg in der politischen Bildung befassen (z. B. Conrad et al., 2022; Weißeno & Grobshäuser, 2021).

In ICCS 2016 war das wahrgenommene Unterrichtsklima das zentrale Instrument, um Unterrichtsqualität zu messen. In 19 von 21 teilnehmenden Ländern und Regionen konnte ein statistisch signifikanter positiver Zusammenhang zwischen einem als offen wahrgenommenen Unterrichtsklima und dem politischen Wissen festgestellt werden (Schulz et al., 2018). Es wird angenommen, dass ein offenes Unterrichtsklima zur eigenen Meinungsbildung ermutigt, ein faires Miteinander ermöglicht und Mitbestimmungsmöglichkeiten bietet (z. B. Lenzi et al.,

2014). Daten einer Studie mit $N = 367$ Schüler*innen legten zudem nahe, dass ein offenes Unterrichtsklima Ungleichheiten im beabsichtigten und tatsächlichen politischen Engagement kompensiert (Weinberg, 2022).

Der Zusammenhang zwischen einem offenen Unterrichtsklima und kognitiven Leistungsmaßen ist jedoch nicht so eindeutig wie die Ergebnisse der ICCS 2016 vermuten lassen. Insbesondere in heterogenen Klassen sorgt ein offenes Unterrichtsklima nicht automatisch für ein fruchtbares Lernumfeld (Munniksmä et al., 2021). Während eine Vielzahl von Studien zu politischer Bildung die Bedeutung des offenen Unterrichtsklimas für den Lernerfolg und die politische Sozialisation der Schüler*innen thematisieren, finden Aspekte wie Klassenführung oder kognitive Aktivierung im Politikunterricht kaum Beachtung (z. B. Eckstein & Noack, 2015; Manganello et al., 2015). Studien, die unterschiedliche Unterrichtsqualitätsmerkmale politischer Bildung gemeinsam betrachten, sind umso seltener (siehe jedoch Bayram-Özdemir et al., 2016). Dementsprechend uneindeutig ist die bisherige Befundlage zur relativen Bedeutsamkeit unterschiedlicher Merkmale insbesondere mit Bezug auf kognitive und nichtkognitive Lernziele.

Die Ergebnisse einer Längsschnittstudie mit $N = 2.012$ Schüler*innen aus Schweden demonstrierten, dass Schüler*innen, die ihre Lehrkraft im Kontext politischer Diskussionen als engagiert wahrnahmen, politischen Themen eher positiv gegenüberstanden (Bayram-Özdemir et al., 2016). Zwischen dem Diskutieren politischer Themen sowie der allgemeinen Schüler*innen-Lehrkräfte-Beziehung und dem Politikempfinden konnte jedoch kein Effekt gefunden werden (Bayram-Özdemir et al., 2016). Hierauf aufbauend konnte anhand von Daten von $N = 594$ österreichischer Schüler*innen gezeigt werden, dass im Kontext politischer Lerngelegenheiten empfundene Emotionen mit der politischen Engagementbereitschaft, der intrinsischen und extrinsischen Lernmotivation sowie dem politischen Wissen zusammenhängen (Graf et al., 2023). Insbesondere Freude, Hoffnung und Stolz hingen positiv mit den nichtkognitiven und kognitiven Lernzielen zusammen (Graf et al., 2023).

1.3.3 Fachfremder Unterricht

Neben der Herausforderung, dass schulische politische Bildung nicht nur im Politikunterricht, sondern auch fächerübergreifend stattfindet und als Schulprinzip in Erscheinung tritt (Achour & Meyer-Heidemann, 2020), ist auch das Fehlen einer einheitlichen Bezeichnung in den Bundesländern und die damit verbundene und bis heute unterentwickelte Identitätsbildung des Faches eine Herausforderung (Pohl, 2021). Vielfach wird die politische Bildung als Teil von Kombinationsfächern gemeinsam mit Geschichte und Erdkunde unterrichtet (Weißeno, 2018).

Zwar bietet die inhaltliche Breite des Faches die Möglichkeit, komplexe politische und gesellschaftliche Inhalte aus unterschiedlichen Perspektiven zu betrachten, jedoch sind Lehrkräfte hierdurch auch mit erhöhten fachwissenschaftlichen Anforderungen konfrontiert (Pohl, 2021).

Unter der Prämisse, dass qualitätsvoller Politikunterricht insbesondere vom Lehrkraftverhalten abhängt (z. B. Bayram-Özdemir et al., 2016), muss angenommen werden, dass der Politikunterricht fachfremder Lehrkräfte qualitativ schlechter ausfällt als der ihrer Kolleg*innen mit entsprechender Ausbildung (Sander, 2020). Das in Anlehnung an das Modell der professionellen Handlungskompetenz von Lehrkräften (Baumert & Kunter, 2006) entwickelte Modell professioneller Handlungskompetenz von Politiklehrer*innen (Weißeno et al., 2013) hält fest, dass Politiklehrkräfte Wissen, Überzeugungen und Motivation benötigen, um im beruflichen Alltag zu bestehen. Trotz individuell variierender Voraussetzungen und Anforderungen gilt, dass professionelle Kompetenz für den Lehrkraftberuf prinzipiell erlern- und vermittelbar ist (Grobshäuser, 2022). Es handelt sich dabei weniger um eine Konsequenz individueller Talente, sondern vielmehr um das Resultat eines gezielten, beruflichen Ausbildungs- und Entwicklungsprozesses (z. B. Kunter et al., 2009).

Es wird angenommen werden, dass fachfremd unterrichtende Politiklehrkräfte im Fachwissen und insbesondere im fachdidaktischen Wissen Defizite haben (Grieger & Oberle, 2020). Empirische Studien, die untersuchen, inwiefern die professionelle Kompetenz von Politiklehrkräften den Zusammenhang zwischen der Lehrkraftausbildung und der Qualität des Politikunterrichts vermitteln, fehlen bislang jedoch (Grobshäuser, 2022; Sander, 2020). Allerdings liegen empirische Ergebnisse zu einzelnen Teilzusammenhängen vor. Eine Befragung von $N = 125$ praktizierenden und angehenden Politiklehrkräften zeigte, dass nur etwa 34,1 Prozent fachfremd unterrichtender Politiklehrkräfte schon einmal vom Beutelsbacher Konsens gehört haben, der für den Politikunterricht jedoch von elementarer Bedeutung ist (Oberle et al., 2018).

Tatsächlich werden politische Bildung sowie weitere Fächer des sozialwissenschaftlichen Spektrums vergleichsweise häufig fachfremd unterrichtet (Grieger, 2022). Die amtliche Schulstatistik in NRW für das Schuljahr 2021/22 belegt, dass das Fach Politik in allen Schulformen und Jahrgangsstufen in NRW so häufig wie kein anderes fachfremd unterrichtet wurde (MKW, 2022).

Eine weitere Herausforderung besteht darin, dass politische Bildung nicht nur intendiert (z. B. im dafür vorgesehenen Leitfach), sondern auch nicht-intendiert in Schulen stattfindet. Ein Beispiel im Kontext des Fachunterrichts sind Mathematikaufgaben zum Thema „Sparen“ oder „Steuern“, welche Reflexionen über soziale Verhältnisse anstoßen können (Del Chicca &

Maaß, 2014). Auch außerhalb des fachunterrichtlichen Settings kann politische Bildung nicht-intendiert stattfinden. So ist die Wahl der*s Klassensprecherin*s nicht nur eine organisatorische Notwendigkeit, sondern auch eine Möglichkeit, demokratische Prozesse einzuüben (Krösche, 2020).

Die Kultusministerkonferenz hält fest, dass das Erlernen und Erfahren von Demokratie eine Querschnittsaufgabe ist und daher alle Lehr- und Fachkräfte in ihrem Unterrichten und Handeln der freiheitlich und rechtsstaatlichen Demokratie verpflichtet sind (Kultusministerkonferenz, 2018). Ziel sei es, junge Menschen dazu zu befähigen, politische, gesellschaftliche und wirtschaftliche Fragen und Probleme kompetent zu beurteilen und aktiv und verantwortungsbewusst für Freiheit, Demokratie, Menschenrechte, Gerechtigkeit, wirtschaftliche Sicherheit und Frieden einzutreten (Kultusministerkonferenz, 2018).

Um zu bestimmen, inwiefern die vorgegebenen Ziele erreicht werden und ob Reformen erfolgreich sind, sind empirische Untersuchungen notwendig. Hierfür muss im Vorfeld festgelegt werden, welche Kompetenzen von den Schüler*innen erwartet werden.

1.4 Bildungsstandards und Kompetenzmodelle für die politische Bildung

Prinzipiell kann im Bildungssystem zwischen input- und outputgeleiteten Kriterien unterschieden werden (Köller, 2020). Das deutsche Bildungssystem war lange Zeit von inputgeleiteten Kriterien (d. h. Haushaltsplänen, Lehrplänen und Rahmenrichtlinien sowie Ausbildungsbestimmungen) geprägt. Im Anschluss an das enttäuschende Abschneiden Deutschlands in der PISA 2000 Studie hat jedoch ein Paradigmenwechsel stattgefunden und outputgeleitete Kriterien rückten in den Fokus (Köller, 2020, McElvany, 2022; für eine kritische Diskussion siehe Sander, 2009). Neben der Vergabe und dem Erwerb von Bildungszertifikaten geht es im Kontext des Outputs von Bildungssystemen vor allem um den Aufbau von Kompetenzen, Qualifikationen, Wissensstrukturen, Einstellungen, Überzeugungen und Werthaltungen (Klieme et al., 2003).

Mit Blick auf Unterricht und Leistungen kann prinzipiell zwischen drei Arten von Outputstandards unterschieden werden (Klieme et al., 2003). Inhaltliche Standards (*content standards*) beschreiben, was Lehrkräfte unterrichten und Schüler*innen lernen sollen. Es werden klar und eindeutig die aufzubauenden Kompetenzen und das zu erreichende Wissen beschrieben. Standards für Lehr- und Lernbedingungen (*opportunity-to-learn standards*) bezeichnen zum Beispiel Schulprogramme sowie das Personal und andere Ressourcen, die Bildungsinstitutionen für einen herausfordernden und gelingenden Unterricht zur Verfügung stehen. Hierunter fällt heutzutage insbesondere auch die digitale Ausstattung von Schulen. Als

Leistungs- oder Ergebnisstandards (*performance standards*) werden Bildungsstandards bezeichnet, die die Lernergebnisse von Schüler*innen zum Gegenstand haben. In den Standards wird beschrieben, über welche Kompetenzen Schüler*innen zu bestimmten Zeitpunkten ihrer Schullaufbahn – meist jedoch am Ende verschiedener Schullaufbahnphasen – verfügen müssen (Klieme et al., 2003). Mithilfe nationaler Tests kann die Einhaltung von Leistungs- und Ergebnisstandards überprüft werden. Die Überprüfung von Schülerleistungen ist ein wichtiger Schritt, damit Bildungsstandards ihr Potential als wichtige Instrumente der Qualitätssicherung und Qualitätsentwicklung von Bildungssystemen entfalten können (Köller, 2020).

Mit der Überprüfung von Kompetenzen mittels standardbezogenen Testverfahren können drei Ziele verfolgt werden (Klieme et al., 2003). Auf der Makro-Ebene kann die systematische Untersuchung von Schülerleistungen dazu dienen, das Erreichen von festgesetzten Bildungsstandards sowie den Erfolg von Bildungsreformen zu kontrollieren (*Bildungsmonitoring*). Dies hilft Bildungsbehörden, Stärken und Schwächen im Bildungssystem zu identifizieren, Trends zu erkennen und gezielte Maßnahmen zur Verbesserung der Bildungsqualität zu ergreifen. Auf der Meso-Ebene können auch einzelne Schulen profitieren, indem sie Testergebnisse nutzen, um den Erfolg von Schulprogrammen oder unterrichtlicher Maßnahmen zu evaluieren (*Schulevaluation*). Schulen bekommen hierdurch zudem die Möglichkeit das Leistungsniveau und die Leistungsheterogenität der eigenen Schüler*innen mit anderen Schulen zu vergleichen, wodurch etwaige Defizite auch auf Schulebene festgestellt werden können. Auf der Mikro-Ebene können Lehrkräfte durch die Überprüfung etwas über die Stärken und Schwächen einzelner Schüler*innen erfahren und gegebenenfalls konkrete Fördermaßnahmen implementieren (*Individualdiagnostik*). Die Kompetenzüberprüfung kann dabei helfen, den individuellen Bedürfnissen der Lernenden gerecht zu werden und den Unterricht adaptiv zu gestalten (Klieme et al., 2003). Außerdem können Kompetenzüberprüfungen den Schüler*innen Feedback zu ihren Leistungen geben. Schüler*innen können Informationen darüber erhalten, was sie bereits gut beherrschen und in welchen Bereichen sie noch Verbesserungsbedarf haben (Klieme et al. 2003). Dieses Feedback ermöglicht es den Lernenden, ihr eigenes Lernen zu reflektieren, Schwächen zu erkennen und gezielt daran zu arbeiten. Insbesondere die Anreicherung der Testergebnisse mit didaktischen Hinweisen kann zur Unterrichtsentwicklung beitragen (Köller, 2020).

1.4.1 Kompetenzmodelle für die politische Bildung in Deutschland

Zu Beginn der 2000er wurden für eine Reihe von Fächern Bildungsstandards verabschiedet und die entsprechend zu erwerbenden Kompetenzen bestimmt (Köller, 2020). Eine der Ausnahmen

ist die schulische politische Bildung, für die es bis heute und trotz unterschiedlicher Initiativen keine verbindlichen nationalen Standards gibt (Weißeno, 2022). Die Initiative zur Bestimmung von Bildungsstandards für ein Unterrichtsfach ging anfänglich vor allem von der Kultusministerkonferenz aus. Eine erste Ausnahme stellte der Entwurf der *Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung* (GPJE) für Standards in der politischen Bildung dar (GPJE, 2004), in dessen Kern die politische Urteils- und Handlungsfähigkeit sowie methodische Fähigkeiten standen. Parallel zum Entwurf der GPJE entstand ein Entwurf für ein Kerncurriculum für den sozialwissenschaftlichen Unterricht in der gymnasialen Oberstufe mit den fünf Kompetenzbereichen (1) Perspektivenübernahme, (2) Konfliktfähigkeit, (3) sozialwissenschaftliches Analysieren, (4) politische Urteilsfähigkeit und (5) Partizipationsfähigkeit (Behrmann et al., 2004). Es stellt sich jedoch die Frage, ob sich die fünf Bereiche der sogenannten *Behrmann-Expertise* nicht den drei Bereichen des GPJE—Modells zuordnen ließen (Sander, 2021b). Zudem wurde kritisiert, dass die Überlegungen auf einem hohen Abstraktionsniveau verbleiben und Konkretisierungen fehlen (Sander, 2021b).

Die verschiedenen Modellvorschläge und insbesondere das GPJE-Modell als verbandspolitischer Vorschlag wurden kritisch in der Fachcommunity diskutiert (Goll, 2019). Zu den Kritikpunkten an den Modellen gehörte, dass die Modelle nur ungenügend theoretisch begründet und empirisch nicht überprüft und nicht anschlussfähig an die Diskussionen der anderen Fachdidaktiken zu der Zeit waren (Weißeno, 2022). Für das Fachwissen wurde zudem bemängelt, dass es im GPJE-Modell als Voraussetzung für die anderen drei Kompetenzbereiche verstanden, jedoch nicht als eigener Kompetenzbereich angesehen wurde. Zudem war zu diesem Zeitpunkt die Bedeutung der Urteilsfähigkeit als domänenspezifischer Kompetenzbereich in Abgrenzung zu einer allgemeinen Kompetenz ungeklärt (Weißeno, 2007).

In der Folge wurde vor allem die Rolle des politischen Wissens thematisiert (z. B. Goll et al., 2011; Juchler, 2010). Einen wichtigen Beitrag für die Weiterentwicklung lieferte das Modell *Konzepte der Politik* von Weißeno et al. (2010), welches das politische Wissen als wichtige Kompetenz anerkennt, den Wissensbegriff für die politische Bildung konkretisiert und fachdidaktisch begründete Basis- und Fachkonzepte identifiziert. Mit dem Modell sollte das Ziel verfolgt werden, Wissensbegriffe zu operationalisieren, die empirisch überprüft und zur Korrektur von Fehlkonzepten eingesetzt werden können (Weißeno et al., 2010).

Kritiker*innen befürchteten eine Verengung der Kompetenzdebatte auf politisches Wissen und bemängelten die Festlegung auf die Politikwissenschaft als einzige Bezugsdisziplin sowie die Auswahl der Konzepte (Autorengruppe Fachdidaktik, 2011). Es wurde zudem befürchtet, dass der strikte Fokus auf die empirische Überprüfbarkeit politischer Kenntnisse dazu

führe, dass die Subjektivität und Plastizität des Lernens außer Acht gelassen werde (Autorengruppe Fachdidaktik, 2011). Die Autor*innen des Modells *Konzepte der Politik* reagierten auf die Kritik mit einem überarbeiteten Kompetenzmodell, dem ein vergleichsweise breiter pädagogisch-psychologischer Kompetenzbegriff zugrunde liegt; das *Modell der Politikkompetenz* von Detjen et al. (2012) integriert das Modell *Konzepte der Politik* in Form von Fachwissen als einen von vier Kompetenzbereichen. Neben dem Fachwissen bilden die politische Urteilsfähigkeit, die politische Handlungsfähigkeit sowie politische Einstellungen und Motivationen die übrigen drei Kompetenzbereiche (Detjen et al., 2012).

Für einige der Bereiche werden reziproke Zusammenhänge angenommen, bspw. zwischen Motivation und Wissen (Detjen et al., 2012). Für andere Bereiche, wie zum Beispiel dem Wissen und der Urteilsfähigkeit, scheint der Zusammenhang jedoch nicht reziprok zu sein, so dass einseitige Abhängigkeiten zwischen den Modellbereichen vermutet werden müssen. Zudem können einige der Kompetenzteilbereiche, insbesondere mit Blick auf die Urteilsfähigkeit und die Handlungsfähigkeit, nur begrenzt in schulischen Lernprozessen gefördert werden (Gloe & Oeftering, 2020). Schließlich liegen nach wie vor nur unzureichende Operationalisierungen der Bereiche Urteilsfähigkeit und Handlungsfähigkeit und folglich keine empirische Überprüfung des Gesamtmodells vor (für eine Übersicht bisheriger empirischer Ergebnisse siehe Weißeno, 2021). Zudem fehlen weiterhin empirische Erkenntnisse zum Verhältnis zwischen unterschiedlichen Kompetenzdimensionen und der Entwicklung von politischer Kompetenz (Sander, 2021b). Bedeutung und Wert des Modells von Detjen et al. (2012) für die Weiterentwicklung der Politikdidaktik als Wissenschaftsdisziplin und für die Stärkung des Kompetenzbegriffes in der politischen Bildung sind jedoch unbestritten (Goll, 2019).

1.4.2 Standards und Assessment Frameworks politischer Bildung im internationalen Kontext

Die Etablierung zentraler Bildungsstandards seit den frühen 2000er Jahren als Grundlage für die Steuerung von Bildungssystemen ist kein allein deutsches Phänomen (Biesta, 2011). Insbesondere die Einführung der *Common Core State Standards* in den USA weist mit Blick auf die Ausgangslage, Motive und Zeitlichkeit einige Ähnlichkeiten zum Prozess in der BRD auf (Hartong, 2015). Eine für die evidenzbasierte Bildungsforschung nachhaltige Wirkung erzielte das im Jahr 2002 in Kraft getretene Gesetzespaket *No Child Left Behind*, welches Zuschüsse zur Bildungsfinanzierung der Bundesstaaten von der Nutzung wissenschaftlich erprobter Methoden und Interventionen abhängig macht (Biesta, 2011). Zwar reicht der Diskurs um Bildungsstandards und Outputsteuerung in den USA sehr viel weiter zurück als in Deutschland, jedoch stellt

die Einführung verbindlicher nationaler Standards für den Mathematik- und Englischunterricht im Jahr 2010 ein Novum für das sonst hochgradig dezentral organisierte Bildungssystem der USA dar (Hartong, 2018).

Verbindliche Standards für das Fach *civic education* existieren nicht, jedoch wurden im Jahr 1994 die freiwilligen *National Standards for Civics and Government* für die Klassenstufen 4, 8 und 12 eingeführt (Center for Civic Education, 1994). Die Standards sind anhand fünf übergeordneter Themenbereiche kategorisiert, in denen es um die Definition von Politik und Regierung, den Grundsätzen des US-amerikanischen politischen Systems, den Aufgaben und Pflichten der US-amerikanischen Regierung, den internationalen Beziehungen der USA sowie der Rolle von Bürger*innen in der US-amerikanischen Demokratie geht (Center for Civic Education, 1994).

Das in den USA durchgeführte *Civics Assessment* des NAEP, welches in regelmäßigen Abständen in der 4., 8. und 12. Klassenstufe durchgeführt wird, orientiert sich stark an den *National Standards for Civics and Government*. Im aktuellen Rahmenkonzept wird neben den Inhaltsbereichen zwischen den Kompetenzdimensionen Wissen, intellektuelle und partizipative Fähigkeiten sowie Civic Dispositionen unterschieden (National Assessment Governing Board, 2018).

Für den Europäischen Kontinent hat der Europarat das Modell der *Competences for Democratic Culture* entwickelt. In dem Modell wird zwischen den Bereichen Werte, Einstellungen, Fähigkeiten sowie Wissen und kritisches Verstehen unterschieden (Council of Europe, 2023). Eine Operationalisierung der Modellbereiche existiert jedoch nicht.

In den vergangenen Jahrzehnten wurden zudem verschiedene internationale LSAs durchgeführt, um die Leistungen und Einstellungen von Jugendlichen im Bereich der politischen Bildung zu messen. Das aktuell wichtigste globale LSA im Kontext der politischen Bildung ist die ICCS. Im Rahmenkonzept der ICCS 2022 wird zwischen einem *Civic Knowledge Framework* bestehend aus Inhalts- und kognitiven Bereichen sowie einem *Civic Attitudes and Engagement Framework* bestehend aus einem Einstellungs- und einem Engagementbereich unterschieden (Schulz et al., 2022).

1.5 Herausforderungen für die empirische Erfassung von Kompetenzen

Anders als Performanz ist Kompetenz nicht unmittelbar beobachtbar. Der Linguist Noam Chomsky differenzierte zwischen Performanz und Kompetenz, indem er Performanz als gezeigtes Verhalten und Kompetenz als die diesem Verhalten zugrundeliegenden Fähigkeiten beschrieb (Chomsky, 1965, S. 4). Nach Chomsky ist Kompetenz einer von mehreren Faktoren

(z. B. Motivation), die Performanz beeinflussen. Dieses Begriffsverständnis wurde zunächst in den USA und später auch international weiterentwickelt und ist ein wichtiger Bestandteil der beruflichen und allgemeinen Bildungspraxis und –forschung (z. B. Vonken, 2017).

International können vor allem die folgenden drei Begriffstraditionen unterschieden werden: eine *behavioristische*, eine *generische* und eine *kognitive* Perspektive (Mulder et al., 2007). Für die behavioristische Perspektive entscheidend ist gezeigtes Verhalten. In der generischen Kompetenzperspektive werden domänenübergreifende Fähigkeiten, die gute Performer von schlechten Performern unterscheiden, in den Blick genommen. Die kognitive Perspektive kommt dem Verständnis von Chomsky am nächsten und stellt kognitive Dispositionen, die eine Handlung ermöglichen, in den Mittelpunkt. Durchgesetzt hat sich vor allem ein von der kognitiven Perspektive eingefärbtes Begriffsverständnis (Stevenson, 1995; Winther, 2020). Messick (1984, S. 28) beschreibt Kompetenz beispielsweise als „what a person knows and can do under ideal circumstances, whereas performance refers to what is actually done under existing circumstances“. Eine im Kontext der deutschen Bildungsforschung gebräuchliche Definition, die an Messicks und Chomskys Verständnis anschließt, beschreibt Kompetenzen als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf bestimmte Klassen von Situationen und Anforderungen beziehen“ (Hartig & Klieme, S. 128; siehe auch Weinert, 2001). Dieses Kompetenzverständnis entspricht einer eher holistischen, das heißt mehrere Bereiche umfassenden, und auf Funktionalität und Erlernbarkeit bezogenen Konzeption von Kompetenz. Nicht unumstritten ist dabei die Frage, inwiefern der Kompetenzbegriff neben kognitiven auch nichtkognitive Komponenten einschließt (z. B. Zlatkin-Troitschanskaia & Seidel, 2011).

Eine Herausforderung für die Beforschung von Kompetenzen besteht darin, die interessierende Kompetenz und ihre Teilbereiche präzise zu konzeptualisieren und zu operationalisieren. Ziel der Konzeptualisierung ist es, ein Konzept in eine klar definierte, theoriebasierte und verständliche Form zu bringen und von anderen Konzepten abzugrenzen (Kühne, 2013). Im Kontext der Operationalisierung werden einem Konzept präzise und beobachtbare Indikatoren zugeordnet, wodurch das Konzept messbar gemacht wird (Kühne, 2013). Im Fall von Kompetenzen steht am Ende dieser beiden Prozesse für gewöhnlich ein Kompetenzmodell. Es kann zwischen Kompetenzstrukturmodellen und Kompetenzniveaumodellen unterschieden werden (Winther, 2020). Während Kompetenzstrukturmodelle insbesondere der Beschreibung und Analyse der dimensional Struktur von Kompetenzen dienen, verfolgen Kompetenzniveaumodelle das Ziel, den Grad an Kompetenz einer Person zu bestimmen und verschiedene Niveaugruppen zu identifizieren (Winther, 2020).

Die Erfassung von Kompetenzmodellen kann mithilfe pädagogisch-psychologischer Tests und Skalen erfolgen. Pädagogisch-psychologische Tests sind wichtige Instrumente für die Qualitätssicherung und Weiterentwicklung eines Bildungssystems, da sie grundlegende und bedeutende Informationsquellen über Einzelpersonen und Gruppen liefern (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014). Gut konstruierte Tests, die für ihre Zwecke valide sind, haben das Potenzial, Bildungsakteur*innen wichtige Erkenntnisse hinsichtlich des Zustands des Bildungssystems und einzelner Teilbereiche zu liefern (AERA et al., 2014). Die ordnungsgemäße Verwendung von Tests kann zu besseren Entscheidungen über Personen, Programme und Reformen führen, als dies ohne ihre Verwendung der Fall wäre. Hierfür unerlässlich ist jedoch, dass einige Mindeststandards beachtet werden.

1.5.1 Qualitätsanforderungen an psychologische Tests

Die traditionellen Hauptgütekriterien für die Entwicklung und Beurteilung von Tests sind Validität, Reliabilität und Objektivität (Moosbrugger & Kelava, 2020). Im internationalen Kontext wird in Ausnahmefällen statt der Objektivität die Testfairness (seltener auch *Equivalence*, Rust & Golombok, 2014) als drittes Hauptgütekriterium angesehen (z. B. AERA et al., 2014; Memon et al., 2010). Weitere Aspekte betreffen die Skalierung, Normierung, Testökonomie, Nützlichkeit, Zumutbarkeit und Unverfälschbarkeit (Moosbrugger & Kelava, 2020).

Die Validität bezieht sich auf das Ausmaß, in der Evidenz und Theorie die Testergebnisinterpretation vor dem Hintergrund des intendierten Verwendungsbereichs eines Tests stützen. Validität wird daher vielfach als Hauptgütekriterium angeführt (Moosbrugger & Kelava, 2020). Wichtig ist hierbei, dass im Kontext der Validierung nicht der Test selbst, sondern die Testannahmen beziehungsweise –interpretationen validiert werden (AERA et al., 2014). Traditionell kann zwischen Inhalts-, Kriteriums- und Konstruktvalidität unterschieden werden (Höfling & Moosbrugger, 2020; siehe jedoch auch AERA et al., 2014). Wichtiger Teil der Testvalidierung ist eine präzise Begriffsbestimmung für die im Rahmen des Tests interessierenden Kenntnisse, Fähigkeiten, Merkmale oder Kompetenzen. Idealerweise wird hierbei für die spätere Überprüfung festgehalten, wie sich das relevante Konstrukt von anderen Konstrukten unterscheiden und verhalten sollte (AERA et al., 2014). Um die Konstruktdeckung entwickelter Tests zu validieren, können Expertenratings zum Einsatz kommen. Die Passung zwischen theoretischer und empirischer Struktur kann mittels Faktorenanalysen überprüft werden.

Das zweite Hauptgütekriterium ist die Reliabilität. In einem allgemeinen Sinn wird Reliabilität als die Konsistenz und Zuverlässigkeit von Messungen oder Ergebnissen in wissenschaftlichen Untersuchungen verstanden (AERA et al., 2014). Eine hohe Reliabilität bedeutet, dass die Messungen zuverlässig sind und nicht (stark) von zufälligen Faktoren oder Messfehlern beeinflusst werden (Moosbrugger & Kelava, 2020). Eine niedrige Reliabilität bedeutet hingegen, dass die Messungen oder Ergebnisse inkonsistent und unzuverlässig sind. Wenn ein Messinstrument nicht zuverlässig ist, kann dies die Genauigkeit der Ergebnisse beeinträchtigen und zu fehlerhaften Schlussfolgerungen führen (Moosbrugger & Kelava, 2020). Daher ist es wichtig, die Reliabilität von Messinstrumenten und Methoden sorgfältig zu überprüfen, um sicherzustellen, dass die Ergebnisse verlässlich sind. Für die Bestimmung der internen Konsistenz einer Skala werden standardmäßig Cronbachs Alpha (α) oder McDonalds Omega (ω) als Reliabilitätsmaß verwendet (Goodboy & Martin, 2020). Cronbachs Alpha ist nur dann geeignet, wenn die Modellannahme der sogenannten tau-Äquivalenz gilt, das heißt jedes Skalenitem in homogener Weise das interessierende Merkmal misst. McDonalds Omega geht hingegen von weniger strikten Modellannahmen aus und ist auch dann geeignet, wenn nicht alle Skalenitems das Merkmal in gleicher Weise messen. Analysen mit simulierten Daten zeigten, dass Cronbachs Alpha die Reliabilität einer Skala tendenziell unterschätzt und McDonalds Omega für große Stichproben zur Überschätzung der Reliabilität neigt (Malkewitz et al., 2023).

Das dritte Hauptgütekriterium ist die Objektivität. Das Gütekriterium der Objektivität hat zum Gegenstand, die Vergleichbarkeit von Testergebnissen von verschiedenen Testpersonen sicherzustellen (Moosbrugger & Kelava, 2020). Ein Test kann als objektiv betrachtet werden, wenn das Testverfahren ausreichend detailliert festgelegt beziehungsweise eingeschränkt ist, damit das Testergebnis unabhängig von Ort, Zeit, Testleiter*in und Auswerter*in ist. Es kann klassischerweise zwischen der Durchführungs- Auswertungs—und Interpretationsobjektivität unterschieden werden (Moosbrugger & Kelava, 2020).

Im internationalen Kontext wird häufig die Testfairness als drittes Hauptgütekriterium angeführt (AERA et al., 2014). Der Begriff Fairness hat in diesem Kontext keine einheitliche technische Bedeutung und wird unterschiedlich verwendet (AERA et al., 2014). Prinzipiell geht es jedoch um die Fairness als grundlegenden Aspekt des Schutzes von Testteilnehmer*innen und der damit einhergehenden Verantwortung von Testentwickler*innen und Testanwender*innen (AERA et al., 2014). Konkret soll sichergestellt werden, dass Testpersonen mit der gleichen Fähigkeit in Bezug auf das Konstrukt, das der Test messen soll, ungeachtet ihrer Zugehörigkeit zu ethnischen, soziokulturellen oder geschlechterspezifischen Gruppen das gleiche

Testergebnis erreichen (Memon et al., 2010). Eine systematische Benachteiligung kann zum Beispielen vorliegen, wenn ein Test, in dem es nicht um sprachliche Kompetenzen geht, hohe Sprachanteile hat, sodass Personen mit geringerer Sprachkompetenz schlechter abschneiden, ohne dass dies an Defiziten in der eigentlich zu messenden Kompetenz liegt. Auch Barrierefreiheit (*barrier-free accessibility*) und universelles Design (*universal design*) sind wichtige Aspekte von Testfairness (AERA et al., 2014). Nach der Durchführung eines Tests können problematische Items mithilfe von *Differential Item Functioning* (DIF) Analysen identifiziert werden (für eine Übersicht zu DIF siehe Penfield, 2016). Ebenfalls geeignet für die Analyse von Subgruppenunterschieden ist die Überprüfung der Messinvarianz von Testinstrumenten (Liu & Dorans, 2016).

1.5.2 Herausforderungen bei der Überprüfung und Interpretation von Testwerten

Eine große Herausforderung für die empirische Erfassung von Kompetenzen ist, dass Kompetenz nicht unmittelbar beobachtbar ist (Chomsky, 1965). Stattdessen erschließt sich das Kompetenzniveau einer Person aus der gezeigten Performanz in einer Vielzahl hinreichend ähnlicher Situationen (Winther, 2020). In der Empirischen Bildungsforschung werden vor allem Tests und Fragebögen eingesetzt, um Kompetenzen zu erfassen (Nagengast & Rose, 2020).

Eine Grundlage für die Frage, wie Merkmalsausprägungen einer Person das Antwortverhalten und damit das Testergebnis einer Person beeinflussen, bilden Testtheorien. Für die Entwicklung und Interpretation psychometrischer Tests kann vor allem zwischen der Klassischen Testtheorie (KTT) und der Probabilistischen Testtheorie beziehungsweise der *Item Response Theory* (IRT) unterschieden werden (Reinders & Post, 2022). Die KTT wird vor allem bei Testitems mit kontinuierlichem Antwortformat eingesetzt (Moosbrugger et al., 2020). Zentrale Annahme der KTT ist, dass eine beobachtete Merkmalsausprägung einer Person Y_i eines Tests i die Summe des wahren Werts τ_i sowie des Messfehlers ε_i ist, das heißt $Y_i = \tau_i + \varepsilon_i$ (Nagengast & Rose, 2020). Es wird weiter angenommen, dass die Merkmalsausprägung Y_i bei wiederholten Messungen konstant bleibt. Der Messfehler ε_i hingegen variiert zufällig um den Wert 0, sodass sich der Mittelwert des Messfehlers mit steigender Anzahl der Messwiederholungen 0 annähert. Die IRT wird hingegen vor allem bei Testitems mit dichotomen Antwortkategorien eingesetzt (Moosbrugger et al., 2020). Die zentrale Annahme der IRT ist, dass die Wahrscheinlichkeit, mit der eine Person ein Item (Frage oder Aufgabe) ‚richtig‘ beantwortet einerseits von der interessierenden Merkmalsausprägung der Person und andererseits von der Schwierigkeit eines Items abhängt (Moosbrugger et al., 2020). IRT-Modell stellen daher einen Zusammenhang zwischen Itemvariablen und Personenvariablen her. Bezogen auf die

Kompetenzerfassung kann mithilfe von IRT die Schwierigkeitsanforderung einer Situation in Beziehung zur Kompetenz einer Person gesetzt werden (Winther, 2020). Ein Vorteil der IRT ist, dass Modellannahmen (z. B. Eindimensionalität und lokale stochastische Unabhängigkeit) und die Güte einzelner Items überprüft werden können (Spoden et al., 2020). Oft unterscheiden sich die Anwendungsfelder von KTT und IRT jedoch, sodass sich die beiden Testtheorien vorteilhaft ergänzen (Moosbrugger et al., 2020).

Eine weitere Herausforderung der empirischen Erfassung von Kompetenzen besteht in der Angemessenheit von Testinhalten für Gruppen mit unterschiedlicher Merkmalsausprägung (Winther, 2020). Im Kontext der Bildungsforschung sind beispielsweise Klassenstufe und Schulform klassische Kategorien entlang derer sich durchschnittliche Kompetenz stark unterscheiden kann. Um derart Unterschiede abbilden zu können, ist es wichtig, dass die einem Kompetenzmodell zugrundeliegenden Indikatoren entwicklungs- beziehungsweise instruktions-sensibel sind (z. B. Messick, 1984). Im Einzelfall ist daher kritisch zu prüfen, inwiefern ein Kompetenzmodell Veränderungen abbilden kann und ob die Binnenstruktur einer Kompetenz gegebenenfalls an sich verändernde Anforderungssituationen in unterschiedlichen Bildungsphasen anzupassen ist (Winther, 2020).

Eine der Prämissen von Kompetenztests ist, dass die *Maximalleistung* von Personen gemessen wird (Asseburg & Frey, 2013; Messick, 1984). Tatsächlich können Testleistungen aber zum Beispiel durch emotionale und motivationale Aspekte wie Testängstlichkeit oder Testmotivation beeinflusst werden (Embse et al., 2018; Spoden et al., 2020). Eine Studie aus Schweden mit $N = 375$ Schüler*innen der 9. Klassenstufe zeigte, dass die Testängstlichkeit, das Interesse an Testinhalten und die Testmotivation statistisch signifikant zwischen unterschiedlichen Testsituation und zwischen verschiedenen Testteilnehmenden variiert (Knekta, 2017). Insbesondere die Testleistungen von weniger fähigen (*low-ability*) Individuen sind von derart ungewünschten Einflussfaktoren betroffen. Außerdem konnte gezeigt werden, dass, in Übereinstimmung mit der Erwartung-Wert-Theorie (*expectancy-value theory*, Wigfield & Eccles, 2000), eine höhere wahrgenommene Testrelevanz mit einer höheren Testmotivation zusammenhängt (Knekta, 2017).

Auch die testspezifische Personenfähigkeit spielt für die Testmotivation eine Rolle. Eine Studie auf Basis der deutschen Teilstichprobe von PISA 2006 mit $N = 9.452$ Schüler*innen demonstrierte, dass die Differenz zwischen Personenfähigkeit und Itemschwierigkeit (*ability-difficulty fit*) positiv mit der Testmotivation zusammenhängt (Asseburg & Frey, 2013). Die Überprüfung von Personen und Item fit, wie sie im Rahmen der IRT entwickelt wurde, kann dazu beitragen, problematische Items zu identifizieren und den Test zu verbessern (Spoden et

al., 2020). Empirische Studien haben gezeigt, dass Schüler*innen eine höhere Testmotivation zeigen, wenn a) Items mental nicht zu herausfordernd sind (z. B. durch die Verwendung von Multiple-Choice-Aufgaben, statt Ausgaben mit offenem Antwortformat), b) die Itemschwierigkeit angemessen ist und c) Schüler*innen ein intrinsisches Interesse am Iteminhalt haben (für einen Überblick siehe Rios, 2021). Zudem kann eine ausreichende Testzeit dazu beitragen, ungewünschte Effekte wie Testängstlichkeit zu reduzieren (Spoden et al., 2020).

Auch der Einsatz computerbasierter Testverfahren kann unter bestimmten Voraussetzungen (z. B. vorhandene Testsoftware, große Stichprobe) sinnvoll sein und zur Qualität von Kompetenzdiagnostiken beitragen (Frey & Hartig, 2013). Beispielsweise können durch den Einsatz computerbasierter Testverfahren neben Ergebnisdaten auch Prozessdaten (z. B. Bearbeitungszeiten und Sequenzen von Bearbeitungsschritten) erfasst werden. Traditionell haben sich Bearbeitungszeiten (*item response times*) als Indikator für unmotiviertes Antwortverhalten bewährt (z. B. Wise, 2017). Die Analyse weiterer Prozessdaten (z. B. nochmaliges Lesen der Aufgabe, Antwortwechsel) hat sich hingegen als wenig hilfreich erwiesen (Welling et al., 2023). Die Identifikation von unmotiviertem Antwortverhalten mittels Antwortzeiten (*rapid-guessing behavior*) und der anschließende Ausschluss von problematischen Antworten kann dazu beitragen, die diagnostische Qualität eines Tests zu erhöhen (Wise, 2017).

1.5.3 Herausforderungen im Kontext der empirischen Erfassung von Civic Literacy

Allgemeine Herausforderungen für die Erfassung von Kompetenzen wie die Qualitätsanforderungen an psychologische Tests sowie die Überprüfung und Interpretation von Testwerten sind auch im Kontext der Forschung zu Civic Literacy von Bedeutung. Hinzu kommen jedoch weitere Herausforderungen, die in besonderem Maße für die Erfassung von Civic Literacy relevant sind.

Civic Literacy ist ein multidimensionales und nicht unmittelbar beobachtbares Konstrukt. Je nach Perspektive schließt es Aspekte unterschiedlicher Wissenschaftsdisziplinen und Unterrichtsfächer ein (z. B. Autorengruppe Fachdidaktik, 2011). Insbesondere der aus dem Englischen entlehnte Civic-Begriff verdeutlicht ein über die reine Politikwissenschaft hinausgehendes Kompetenzverständnis. Das damit einhergehende inhaltliche wie auch kompetenztheoretisch breite pädagogisch-psychologische Begriffsverständnis erfordert eine präzise Definition und Operationalisierung der unterschiedlichen Kompetenzbereiche, die empirische Überprüfung der im Zuge der Modellentwicklung getroffenen Annahmen sowie die sorgfältige Interpretation der Testergebnisse und der sich hieran gegebenenfalls anschließenden Modellüberarbeitung.

Ein Schwerpunkt bei der Entwicklung und Überprüfung des Civic Literacy Kompetenzmodells ist die Zusammenhangsstruktur (d. h. Einteilung in Dimensionen und Facetten). Die aus der theoretischen Definition des Civic Literacy Modells abgeleiteten Faktoren können mittels konfirmatorischer Faktorenanalysen inferenzstatistisch überprüft werden (Gäde et al., 2020). Konfirmatorische Faktorenanalysen ermöglichen die Überprüfung theoretisch abgeleiteter Hypothesen über die faktorielle Struktur eines Variablensets und den Vergleich alternativer Modelle (Landwehr, 2017). Die Klärung der faktoriellen Struktur des Civic Literacy Modells ist dabei nicht nur aus theoretischer Sicht, sondern auch aus diagnostischer Sicht relevant. Die Abwägung ob Gesamtestwerte oder Testwerte einzelner Testbereiche berichtet werden sollten, wird maßgeblich durch die empirisch begründete Zuordnung von Items zu einzelnen Faktoren bestimmt (Gäde et al., 2020). Ein Beispiel im Kontext von Civic Literacy bildet die politische beziehungsweise zivilgesellschaftliche Partizipation und Partizipationsbereitschaft. Während aus theoretischer Sicht eine Unterscheidung zwischen der politischen und zivilgesellschaftlichen Partizipation sinnvoll erscheint (Ekman & Amnå, 2012), steht eine empirische Überprüfung dieser Annahme bislang aus. Auch allgemein liegen bislang nur für einzelne Teilbereiche empirische Überprüfungen der faktoriellen Struktur von Civic Literacy beziehungsweise politischer Kompetenz mittels konfirmatorischer Faktorenanalyse vor (z. B. Bohm, 2015; Velez & Knowles, 2022).

Ein weiterer Schwerpunkt der Überprüfung des Civic Literacy Modells ist die Frage nach schwierigkeitsgenerierenden Itemmerkmalen im Kontext des politischen Wissenstests. Ziel des politischen Wissenstests ist es, das politische Wissen von Schüler*innen unterschiedlicher Klassenstufen zu erfassen. Um zwischen unterschiedlichen Merkmalsausprägungen unterscheiden zu können, ist es wichtig, dass die Items von Personen mit unterschiedlicher Merkmalsausprägung auch unterschiedliche Antworten hervorrufen. Entsprechend sollten extrem leichte und extrem schwierige Items vermieden beziehungsweise sparsam eingesetzt werden (Moosbrugger & Brandt, 2020). Die Zielschwierigkeit der Testitems ist in Abhängigkeit zur erwarteten Merkmalsausprägung in der interessierenden Gesamtpopulation zu wählen. Sofern die Merkmalsausprägung in der Zielpopulation unbekannt sein sollte, kann ähnlich wie bei der Identifikation von Schwellenwerten beziehungsweise Kompetenzniveaus (Brückner et al., 2020) auf Einschätzungen von Expert*innen zurückgegriffen werden. Beeinflusst werden Itemschwierigkeiten durch verschiedene Aufgabenmerkmale. Für die Itemschwierigkeit relevante Aufgabenmerkmale beziehen sich auf Anforderungen, die beim Bearbeiten und Lösen von Testitems bewältigt werden müssen (Rauch & Hartig, 2020).

Neben der Struktur und der Komplexität von Aufgaben ist auch der Grad der möglichen Exposition mit den Aufgabeninhalten von Bedeutung. Der Grad der möglichen Exposition beschreibt das allgemeine Vorhandensein von Informationen und die Wahrscheinlichkeit, dass ein*e Testteilnehmer*in mit dieser Information bereits in Kontakt gekommen ist (Juhasz et al., 2019; Stanovich & Cunningham, 1993). Beispielsweise sind Wortschatztests mit weniger frequentierten Wörtern schwieriger zu lösen als Wortschatztests mit stark frequentierten Wörtern (z. B. Brysbaert et al., 2018). Neben der durchschnittlichen Worthäufigkeit kann auch das Sprachsetting, in dem die Wörter eines Items vorkommen, die Schwierigkeit eines Items erklären. Für das Sprachsetting kann zwischen alltäglichen und bildungssprachlichen Sprachsettings unterschieden werden (Brysbaert et al., 2011). Auch im Kontext politischen Wissens sind Anforderungen bei der Anwendung der Fachsprache eine wichtige Möglichkeit zur Steuerung der Itemschwierigkeit (GPJE, 2004). Insbesondere aufgrund des sprachintensiven Lernens im Kontext politischer Bildung (Drippe, 2018) kann die Untersuchung von Wortschatzhäufigkeiten in unterschiedlichen Sprachsettings einen Beitrag zur Aufklärung von Itemschwierigkeiten leisten.

Ein letzter Schwerpunkt ist das Verhältnis von Selbsteinschätzungen und objektiv gemessener Performanz im Kontext des politischen Wissenstests. Die Wahrnehmung der eigenen Performanz kann durch Selbsteinschätzungen ausgedrückt werden. Selbsteinschätzungen sind ein wichtiges Instrument zur Entscheidungsfindung in alltäglichen Situationen (Freund & Kasten, 2012). Insbesondere in Situationen, in denen eine objektive Bewertung von eigenen Fähigkeiten aus Zeit- oder anderen strukturellen Gründen nicht möglich ist, sind Selbsteinschätzungen häufig die einzige beziehungsweise wichtigste Information, um die Erfolgserwartung einer bestimmten Handlung abschätzen zu können. Im Verbund mit objektiven Leistungsüberprüfungen können Selbsteinschätzungen auf unterschiedliche Weise (z. B. absolut oder relativ zu Peers) und zu unterschiedlichen Zeitpunkten (z. B. vor oder nach einer objektiven Leistungsüberprüfung) erfolgen (Ackerman & Wolman, 2007). Aus forschungspraktischer Perspektive sind Selbsteinschätzungen ein hilfreiches Instrument, um Merkmalsausprägungen in forschungswirtschaftlicher Weise zu erfassen (Klinger, 2022). Aus einer Bildungsperspektive sind Selbsteinschätzungen zudem ein wichtiger Faktor für den Lern- und Bildungserfolg (z. B. Vuorre & Metcalfe, 2022), indem Lernende mittels Selbsteinschätzungen Lerndefizite erkennen und ihr Lernen anpassen können (Dunlosky & Rawson, 2012). Es bestehen jedoch systematische Unterschiede zwischen Selbsteinschätzungen und tatsächlicher Performanz. Eine Meta-Studie auf Basis von 41 Studien zeigte, dass Selbsteinschätzungen und Performanz nur moderat miteinander zusammenhängen ($r = .33$; Freund & Kasten, 2012). Ein Befundmuster,

welches die systematischen Unterschiede zwischen Selbsteinschätzung und Performanz beschreibt, ist der Dunning Kruger-Effekt (Kruger & Dunning, 1999). Der Dunning Kruger-Effekt beschreibt, dass sich leistungsschwache Individuen deutlich überschätzen, wohingegen sich leistungsstarke Individuen leicht unterschätzen. Grund hierfür sei, dass leistungsschwachen Individuen die metakognitiven Fähigkeiten fehlten, die notwendig wären, um die eigene Inkompetenz zu erkennen (Kruger & Dunning, 1999). Der Befund konnte auch für das politische Wissen repliziert werden (Anson, 2018; Rapeli, 2023). Zudem wurde festgestellt, dass Individuen, die ihr politisches Wissen überschätzten, häufiger zu radikalen politischen Einstellungen neigten (Van Prooijen et al., 2018) und eher populistische Parteien wählten (Van Prooijen, 2021). Umstritten, ist jedoch die Annahme, dass ein Defizit in metakognitiven Fähigkeiten für das Befundmuster des Dunning-Kruger Effekts verantwortlich ist (z. B. Krueger & Mueller, 2002). Kritiker*innen argumentieren, dass der Dunning-Kruger Effekt ein statistisches Artefakt sei, welches auf einen better-than-average Effekt und eine regression-toward-the-mean zurückzuführen sei (Gignac & Zajenkowski, 2020). Außerdem ist bisher unklar, inwiefern das Verhalten von Lehrkräften eine Rolle für den Dunning-Kruger Effekt spielt.

1.6 Forschungsfragen

Es besteht weitgehend Einigkeit darüber, dass Civic Literacy neben anderen Basiskompetenzen ein wichtiges Bildungsziel der Sekundarstufe I ist (z. B. Aktionsrat Bildung, 2020; Detjen et al., 2012; Schulz et al., 2018). Ähnlich zu den in anderen Fachdisziplinen geführten Kompetenzdebatten wurde auch im Kontext der Politikdidaktik darüber diskutiert, was schulische politische Bildung zu leisten hat und welche Bereiche ein Kompetenzmodell abdecken sollte (Goll, 2019). Trotz unterschiedlicher Entwürfe aus dem US-amerikanischen (z. B. National Assessment Governing Board, 2018) und dem deutschen Kontext (z. B. Abs & Hahn-Laudenberg, 2017; Detjen et al. 2012) sowie aus international vergleichender Perspektive (Schulz et al., 2016) bestehen einige Desiderate weiter fort. Beispielsweise fehlt ein auf den deutschen Kontext angepasstes Kompetenzmodell für die Sekundarstufe I, das auf fachdidaktische Überlegungen aufbaut, vollständig operationalisiert ist und sich für den Einsatz in LSAs eignet (Oberle & Weißeno, 2017). Zudem wurden bisherige Modelle für querschnittliche Erhebungen konzipiert. Wünschenswert wäre jedoch ein Kompetenzmodell, welches eine längsschnittliche Betrachtung von Civic Literacy in der Adoleszenz ermöglicht, um wichtige Erkenntnisse zu Bedingungsbeziehungen und Entwicklungsdynamiken zu erhalten (Kuger & Gille, 2020).

Eine besondere Rolle für die Entwicklung und Manifestation von Civic Literacy kommt der Schule und insbesondere dem Politikunterricht zu (z. B. Castillo et al., 2015; Martens & Gainous, 2013). Nichtsdestotrotz fehlen auch hier bislang wichtige Erkenntnisse, vor allem hinsichtlich der Bedeutung verschiedener Unterrichtspraktiken und –methoden mit Blick auf die verschiedenen Teilbereiche von Civic Literacy (siehe jedoch Bayram-Özdemir et al., 2016 sowie Maurissen, 2020). Schließlich gilt es, die Beforschung von Civic Literacy mit aktuellen Diskursen in der pädagogischen Psychologie zu verknüpfen. Ein Beispiel hierfür ist die Frage, inwiefern subjektiv empfundene und objektiv gemessene Fähigkeiten zusammenhängen (z. B. Hofer et al., 2022).

Vor diesem Hintergrund wurde in der vorliegenden Dissertation untersucht, wie Civic Literacy bei Schüler*innen der Sekundarstufe I gemessen werden kann, welche Unterrichtsmerkmale mit verschiedenen Zielen schulischer politischer Bildung zusammenhängen und inwiefern der Dunning-Kruger Effekt als weithin bekanntes pädagogisch-psychologisches Phänomen auch im Kontext der politischen Bildung gefunden und analysiert werden kann. Hierfür wurden die folgenden übergreifenden Forschungsfragen untersucht:

- (1) Was umfasst Civic Literacy und wie kann sie bei Jugendlichen gemessen werden?
 - a. Auf Basis welcher Dimensionen und Facetten kann Civic Literacy konzeptualisiert und operationalisiert werden?
 - b. Werden die Dimensionen und Facetten des Civic Literacy Kompetenzmodells durch empirische Daten gestützt?
 - c. Inwiefern werden Itemschwierigkeiten in einem politischen Wissenstest von Wortfrequenzen beeinflusst?
- (2) Inwiefern stehen unterschiedliche Merkmale der schulischen politischen Bildung mit den verschiedenen Bereichen von Civic Literacy in Zusammenhang?
 - a. Wie hängen die Unterrichtsqualitätsmerkmale (a) offenes Unterrichtsklima, (b) kognitive Aktivierung und (c) das Diskutieren aktueller politischer Themen mit Partizipationsbereitschaft zusammen?
 - b. Inwiefern werden die Zusammenhänge durch (a) politisches Interesse und (b) politisches Wissen mediiert?
- (3) Wie hängen subjektives und objektives politisches Wissen zusammen?
 - a. Gibt es Unterschiede im Zusammenhang von subjektivem und objektivem politischem Wissen zwischen der 7. und 10. Klassenstufe?
 - b. Inwiefern hängen Lehrkraftverhalten und Persönlichkeitsmerkmale mit der Differenz zwischen subjektivem und objektivem politischem Wissen zusammen?

1.7 Zusammenfassungen der Einzelbeiträge

Die kumulative Dissertation besteht aus einer theoretischen Arbeit und drei empirischen Studien (siehe Tabelle 1a) sowie einem Beitrag mit weiterführenden Analysen (siehe Tabelle 1b). Für die empirischen Beiträge der Dissertation (*Beitrag II, III, IV und V*) wurden Daten der *Studie zur Entwicklung politischer und gesellschaftlicher Kompetenz im Jugendalter* (EPKO) verwendet. Die folgenden Ergebniszusammenfassungen sind gleichermaßen strukturiert und untergliedern sich in die Abschnitte *Theoretischer Hintergrund, Forschungsziele, Methode, Ergebnisse* und *Schlussfolgerung*.

1.7.1 *Beitrag I: Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung*

Theoretischer Hintergrund. Die Kompetenzdebatte der 2000er Jahre hat auch im Kontext der schulischen politischen Bildung für eine Diskussion zur Notwendigkeit von Bildungsstandards (Sander, 2009) und der Frage, welche Kompetenzen gelernt werden sollen, geführt (Goll, 2019). Ähnlich zu den Entwicklungen in anderen Fächern, wurden auch für die politische Bildung unterschiedliche Kompetenzmodelle entwickelt (z. B. Detjen et al. 2012; GPJE, 2004), mit dem Ziel, politische und gesellschaftliche Kompetenz (Civic Literacy) zu messen. Die konzeptualisierten Modelle unterschieden vor allem zwischen politischem Wissen einerseits und weiteren Handlungsdispositionen in den Bereichen Motivation, Einstellung und Volition andererseits (z. B. Autorengruppe, 2011; Detjen et al., 2012; Weißeno et al., 2010). Nichtsdestotrotz fehlt weiterhin ein fachdidaktisch begründetes und vollständig operationalisiertes Kompetenzmodell, welches die systematische Erfassung von Civic Literacy in der Sekundarstufe I ermöglicht.

Forschungsziel. Ziel des Beitrags ist die Entwicklung eines Civic Literacy Kompetenzmodells für die Sekundarstufe I, welches in LSAs sowie in Längsschnittstudien zum Einsatz kommen kann. Hierfür werden einzelne Kompetenzbereiche identifiziert und deren Bedeutung für das Gesamtmodell beschrieben. Zudem werden die dem entwickelten Wissenstest zugrundeliegenden Items hinsichtlich ihrer Relevanz und ihrer inhaltlichen Verordung im Modell durch ein Expertengremium bewertet.

Methode. Die Entwicklung des Kompetenzmodells basiert auf existierenden Kompetenzmodellen, empirischen Untersuchungen und theoretischen Überlegungen im Kontext der politischen Bildung. Auf der Grundlage dieser Erkenntnisse werden die unterschiedlichen Bereiche des Kompetenzmodells destilliert. Schließlich werden die Itembewertungen der

Mitglieder des Expertengremiums mit Blick auf die Übereinstimmung der zuvor vorgenommenen Verortung im Wissenstest analysiert.

Ergebnisse. Es wurde ein Civic Literacy Kompetenzmodell mit den Bereichen politisches Wissen, politische Motivationen, politische Einstellungen und politische Volitionen entwickelt. Für jeden der Bereiche wurde zudem eine Binnenstruktur identifiziert. Der Bereich des politischen Wissens besteht aus drei Teilfähigkeiten und drei Inhaltsbereichen. Die übrigen Kompetenzmodellbereiche (Motivation, Einstellung, Volition) bestehen aus jeweils zwei Facetten. In Bezug auf den Wissenstest zeigte eine Analyse des Expertenratings zudem, dass die jeweilige Teilfähigkeit beziehungsweise der jeweilige Inhaltsbereich für 78 % aller Testitems der siebten und für 74 % aller Testitems der zehnten Klassenstufe richtig geratet wurde (d. h. das Rating stimmte mit der zuvor von den Itementwickler*innen vorgenommenen Einteilung der Items überein).

Schlussfolgerung. Das entwickelte Kompetenzmodell mit seinen vier Bereichen kann nach einer empirischen Erprobung zukünftig in Studien der politischen Kultur- und Sozialisationsforschung sowie der Empirischen Bildungsforschung und der politischen Psychologie Anwendung finden. Das Expertenrating demonstriert zudem, dass der politische Wissenstest unterschiedliche Inhaltsbereiche und Teilfähigkeiten in ausgewogenem und nachvollziehbarem Maße abbildet.

1.7.2 Beitrag II: Towards the Measurability of a Competence Model for School-based Civic Education

Theoretischer Hintergrund. Damit psychologische Tests ihr Potential entfalten können, Bildungsakteur*innen wichtige Erkenntnisse hinsichtlich des Zustands eines Bildungssystems und einzelner Teilbereiche zu liefern, muss gewährleistet sein, dass die Tests bestimmten Gütekriterien entsprechen (AERA et al., 2014). Für Kompetenzmodelle bedeutsam ist, dass die theoretisch angenommene Binnenstruktur durch empirische Daten gestützt wird, dass das Konstrukt in ähnlicher Weise in unterschiedlichen Gruppen gemessen werden kann, dass im Fall von Längsschnitterhebungen das Modell entwicklungs- beziehungsweise alterssensibel ist sowie die Frage, wie sich die Messung eines Konstrukts zu Messungen eines ähnlichen (konvergente Validität) beziehungsweise eines unähnlichen (diskriminante Validität) Konstrukts verhält (Messick, 1984; Moosbrugger & Höfling, 2020).

Forschungsziel. Ziel des Beitrags ist es, auf Basis empirischer Daten die Gültigkeit der im Kontext der Entwicklung des Civic Literacy Kompetenzmodells getroffenen Annahmen zu überprüfen und wo immer notwendig entsprechende Anpassungen vorzunehmen.

Methoden. Die empirische Untersuchung basiert auf Daten von insgesamt 1.047 Schüler*innen der 7. ($n = 613$; $M_{\text{Alter}} = 12,47$; 47,4 % weiblich; 44,4 % Migrationshintergrund) und 10. Klassenstufe ($n = 434$; $M_{\text{Alter}} = 15,49$; 47,7 % weiblich; 35,9 % Migrationshintergrund). Die Daten wurden im Klassenkontext erfasst und stammen von insgesamt 69 Schulklassen aus NRW. Für die Beantwortung der Forschungsfragen zur Binnenstruktur wurden konfirmatorische Faktorenanalysen (KFA) gerechnet. Die Messinvarianz wurde mithilfe von Mehr-Gruppen KFA ermittelt. Mittelwertunterschiede zwischen der 7. und 10. Klassenstufe wurden mittels Zweistichproben-t-tests für unabhängige Stichproben analysiert. Für die konvergente und diskriminante Validität der Konstrukte wurde ein nomologisches Netzwerk mit drei Ebenen identifiziert und anhand der Korrelationen zwischen den Konstrukten untersucht.

Ergebnisse. Die theoretisch angenommene Struktur wurde weitgehend von den empirischen Daten gestützt. Ausnahmen waren das Effektivitätsgefühl als Teil der Dimension politische Motivation, für das sich eine Unterscheidung in internes und externes Effektivitätsgefühl zeigte. Außerdem wurde die Unterscheidung zwischen politischer und zivilgesellschaftlicher Partizipationsbereitschaft als Teil der Dimension politische Volition nicht durch die KFA gestützt. Die Mehr-Gruppen KFA haben ergeben, dass hinsichtlich der 7. und 10. Klassenstufe skalare Messinvarianz für das Kompetenzmodell angenommen werden kann. Die Mittelwerte aller Facetten des Kompetenzmodells waren in der 10. Klassenstufe höher als in der 7. Klassenstufe. Auch die Ergebnisse der Untersuchung zur konvergenten und diskriminanten Validität der Konstrukte entsprachen weitgehend den vorher getroffenen Annahmen. Einzige Ausnahme war die Einstellung zur pluralen Gesellschaft (Dimension politische Einstellung), die weniger stark mit dem politischen Interesse (Dimension politische Motivation) als mit dem Gefühl sozialer Zugehörigkeit (theoretisch nicht verbundenes Konstrukt) korrelierte. Der Unterschied in den Korrelationen war jedoch nicht statistisch signifikant.

Schlussfolgerung. Die Modellannahmen werden weitgehend von den empirischen Daten gestützt. Auf Basis der Analysen wurden kleine Veränderungen an der Kompetenzmodellstruktur vorgenommen. Außerdem legen die Analysen nahe, dass das Kompetenzmodell in einer längsschnittlichen Studie eingesetzt werden kann.

1.7.3 Beitrag III: What Makes Domain Knowledge Difficult? Word Usage Frequency from SUBTLEX and dlexDB Explains Knowledge Item Difficulty

Theoretischer Hintergrund. In der pädagogischen Psychologie sind Wissenstests wichtige Bestandteile von Kompetenzüberprüfungen und dienen vor allem dazu, den Lernerfolg von

Schüler*innen zu verfolgen (National Research Council, 2012). Mit einem Wissenstest soll die interessierende Merkmalsausprägung einer Person gemessen werden. Die Itemschwierigkeit soll dabei ausdrücken, wie schwierig ein Item im Sinne des zu erfassenden Konstrukts ist (Embretson & Reise, 2013). Tatsächlich kann die Itemschwierigkeit aber durch externe Störfaktoren wie Testmotivation und Testängstlichkeit beeinflusst werden (z. B. Spoden et al., 2020). Eine wichtige Rolle im Kontext politischen Wissens könnte dabei spielen, inwiefern Lernende die Möglichkeit hatten, mit den in einem Item vorkommenden Konzepten in Kontakt zu kommen (siehe *environmental opportunity hypothesis*: Stanovich & Cunningham, 1993). Ein hilfreicher Proxy zur Bestimmung des potentiellen Expositionsgrads eines Items sind Worthäufigkeiten in alltäglichen beziehungsweise akademischen Sprachsettings (Brysbaert et al., 2018).

Forschungsziel. Ziel des Beitrags ist es zu ermitteln, inwiefern die durchschnittliche Häufigkeit von Wörtern eines Items in alltäglichen und akademischen Sprachsettings mit der Schwierigkeit eines Items zusammenhängt. Konkret wird untersucht, inwiefern die beiden Worthäufigkeiten separat in unterschiedlichen Modellen (alltäglich vs. akademisch) und gemeinsam in einem Modell (alltäglich und akademisch) mit der Itemschwierigkeit zusammenhängen.

Methoden. Die Analysen der Studie basieren auf Daten von insgesamt 250 Schüler*innen der 7. ($n = 152$; $M_{\text{Alter}} = 12,54$; 44,8 % weiblich; 30,8 % Migrationshintergrund) und 10. Klassenstufe ($n = 98$; $M_{\text{Alter}} = 16,12$; 35 % weiblich; 32,7 % Migrationshintergrund). Die Daten wurden im Klassenkontext erfasst und stammen von insgesamt 17 Schulklassen aus NRW und Thüringen. Um den Zusammenhang zwischen den beiden Indikatoren der Worthäufigkeit und der Itemschwierigkeit zu erklären, wurden drei *Ordinary Least Squares* Regressionen (OLS) gerechnet, die (1) nur den Indikator für die alltägliche, (2) nur den Indikator für die akademische und (3) beide Indikatoren der Worthäufigkeit als erklärende Variablen enthalten.

Ergebnisse. Eine hohe Worthäufigkeit in alltäglichen Sprachsettings hing statistisch signifikant negativ mit der Itemschwierigkeit zusammen. Konkret heißt das, dass Items mit Wörtern, die häufig in alltäglichen Sprachsettings verwendet werden, tendenziell leichter waren, als Items mit Wörtern, die weniger häufig in alltäglichen Sprechsettings verwendet werden. Zwischen der Worthäufigkeit in akademischen Sprachsettings und der Itemschwierigkeit bestand kein statistisch signifikanter Zusammenhang. Wenn beide Indikatoren der Worthäufigkeit in einem gemeinsamen Modell untersucht wurden, hing die alltägliche Worthäufigkeit weiter statistisch signifikant negativ mit der Itemschwierigkeit zusammen. Die akademische Worthäufigkeit hing in diesem Modell statistisch signifikant positiv mit der Itemschwierigkeit

zusammen. Die Koeffizienten beider Indikatoren waren in dem gemeinsamen Modell (3) statistisch signifikant größer als in den Modellen (1) und (2), in denen jeweils nur ein Indikator inkludiert war.

Schlussfolgerung. Die durchschnittliche Worthäufigkeit in Items hängt statistisch signifikant mit der Schwierigkeit von Items zusammen. Unklar bleibt jedoch, inwiefern dieser Zusammenhang auf den Expositionsgrad von in Items behandelten Konzepten oder Aspekten der Lesbarkeit von Wörtern (z. B. Komplexität und Länge) zurückzuführen ist. Ein spannender Befund ist zudem, dass relative Unterschiede zwischen Worthäufigkeiten in alltäglichen und akademischen Sprachsettings zusätzliche Varianz in der Itemschwierigkeit aufklären.

1.7.4 Beitrag IV: Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators

Theoretischer Hintergrund. Die schulische politische Bildung ist ein wichtiges Instrument, um Civic Literacy zu vermitteln (Galston, 2001; Kultusministerkonferenz, 2018). Die Förderung der politischen Partizipation und Partizipationsbereitschaft erfolgt im Politikunterricht indirekt über die Vermittlung anderer Dispositionen wie dem politischen Interesse oder das politische Wissen (Oberle & Schwanholz, 2023). Eine prinzipiell wichtige Rolle bei der Erreichung von Lernzielen spielt die Unterrichtsqualität (Klieme, 2022). Der Diskurs zu fachspezifischen Aspekten der Unterrichtsqualität ist im Kontext der politischen Bildung bisher jedoch nur cursorisch geführt worden (siehe jedoch Conrad et al., 2022; Weißeno & Grobshäuser, 2021). Insbesondere der Vergleich der Indikatoren offenes Unterrichtsklima (z. B. Eckstein & Noack, 2015), kognitive Aktivierung (Weißeno & Landwehr, 2015) und das Diskutieren aktueller politischer Themen (Bittman & Russel, 2016) könnte für die politische Bildung interessant sein. Bislang liegen jedoch kaum empirische Studien vor, die einen Vergleich unterschiedlicher Unterrichtsqualitätsmerkmale und deren Bedeutung für die Zielerreichung in der politischen Bildung zulassen (siehe jedoch Bayram-Özdemir et al., 2016).

Forschungsziel. Ziel des Beitrags ist es, die relative Bedeutung unterschiedlicher Unterrichtsqualitätsmerkmale für die politische und zivilgesellschaftliche Partizipationsbereitschaft zu ermitteln. Außerdem wird untersucht, inwiefern der Zusammenhang durch politisches Wissen und politisches Interesse mediiert wird.

Methoden. Die Analysen der Studie basieren auf den Daten von insgesamt 250 Schüler*innen der 7. ($n = 152$: $M_{\text{Alter}} = 12,54$; 44,8 % weiblich; 30,8 % Migrationshintergrund) und 10. Klassenstufe ($n = 98$: $M_{\text{Alter}} = 16,12$; 35 % weiblich; 32,7 % Migrationshintergrund). Die Daten wurden im Klassenkontext erfasst und stammen von insgesamt 17 Schulklassen aus

NRW und Thüringen. In einem ersten Strukturgleichungsmodell wurde der Zusammenhang zwischen offenem Unterrichtsklima, kognitiver Aktivierung sowie dem Diskutieren aktueller politischer Themen und Partizipationsbereitschaft untersucht. In einem zweiten Strukturgleichungsmodell wurde überprüft, inwiefern die Zusammenhänge durch politisches Interesse und politisches Wissen mediiert werden.

Ergebnisse. Im Modell ohne Mediatoren hingen das offene Unterrichtsklima und kognitive Aktivierung positiv mit der Partizipationsbereitschaft von Schüler*innen zusammen. Für das Diskutieren aktueller politischer Themen wurde kein statistisch signifikanter Zusammenhang gefunden. Im Modell mit den Mediatoren hing das offene Unterrichtsklima weiterhin direkt mit Partizipationsbereitschaft zusammen. Zudem bestand ein negativer Zusammenhang mit dem politischen Wissen. Der Zusammenhang zwischen der kognitiven Aktivierung und der Partizipationsbereitschaft wurde vollständig über das politische Interesse und politische Wissen mediiert. Das Diskutieren aktueller politischer Themen hing positiv mit dem politischen Interesse zusammen. Eine separate Analyse der Daten der 7. und 10. Klassenstufe bestätigte das Befundmuster weitgehend.

Schlussfolgerung. Verschiedene Unterrichtsqualitätsmerkmale können je nach Lernziel eine unterschiedliche Bedeutung haben. Daher ist eine gemeinsame Betrachtung unter Berücksichtigung der relativen Bedeutung der verschiedenen Merkmale wichtig. Insbesondere interessant ist, dass das sonst als überaus wichtig erachtete offene Unterrichtsklima negativ mit dem politischen Wissen zusammenhängt. Ein Grund hierfür könnte sein, dass ein sehr offenes Unterrichtsklima den regulären Lernprozess und somit die Vermittlung von Wissen stört.

1.7.5 Beitrag V (weiterführende Analysen): When Will They Know What They Don't Know? – Political knowledge and the Infamous Dunning-Kruger Effect

Theoretischer Hintergrund. Die akkurate Selbsteinschätzung der eigenen Fähigkeit ist in einer Vielzahl von Lebenssituationen von Bedeutung. Im Bildungskontext sind akkurate Selbsteinschätzungen mit Wohlbefinden, höherer Motivation sowie Lern- und Bildungserfolg assoziiert (Bouffard et al., 2011; Sheldrake et al., 2022). Im Kontext politischer Bildung wurde zudem gefunden, dass Personen, die ihr politisches Wissen überschätzen, eher zu radikalen politischen Ansichten neigen (z. B. van Prooijen, 2021). Empirische Untersuchungen zeigen jedoch, dass sich Selbsteinschätzungen und tatsächliche Testleistungen systematisch unterscheiden (z. B. Zell & Krizan, 2014). Auf Basis einer Befragung von Studierenden der Cornell Universität fanden Justin Kruger und Daniel Dunning (1999), dass leistungsschwache Studierende zu einer Überschätzung und leistungsstarke Studierende zu einer Unterschätzung neigten,

wobei die leistungsstarken Studierenden sich prinzipiell besser einschätzten als die leistungsschwachen Studierenden. Dieses Befundmuster konnte vielfach repliziert werden und ist als sogenannter Dunning-Kruger Effekt bekannt geworden. Die Begründer des Dunning-Kruger Effekts sowie weitere Wissenschaftler*innen (z. B. Hiller, 2023; Schlösser et al., 2013; Zhou & Jenkins, 2020) gehen davon aus, dass leistungsschwache Individuen nicht die metakognitiven Fähigkeiten haben, um ihr Leistungsdefizit erkennen zu können (Dunning, 2011). Kritiker*innen halten dem entgegen, dass der Dunning-Kruger Effekt ein statistisches Artefakt und auf den better-than-average Effekt und regression-to-the-mean zurückzuführen sei (z. B. Gignac & Zajenkowski, 2023; Magnus & Peresetsky, 2022). Insbesondere im Schulkontext spielt das Erlernen metakognitiver Fähigkeiten und die damit verbundene Fähigkeit, eigene Lernfortschritte beurteilen und Lerndefizite erkennen zu können, eine große Rolle (z. B. Ohtani & Hisasaka, 2018; Perry et al., 2019). Eine wichtige Position nehmen dabei die Lehrkräfte ein, die das Erlernen metakognitiver Fähigkeiten auf unterschiedliche Art und Weise fördern können (z. B. Branigan & Donaldson, 2020; Dignath & Büttner, 2008).

Forschungsziel. Ziel des Beitrags ist es, den Dunning-Kruger Effekt in Bezug auf das politische Wissen von Schüler*innen der Sekundarstufe I zu untersuchen. Konkret wird untersucht, inwiefern sich der Dunning-Kruger Effekt in den Klassenstufen 7 und 10 unterscheidet. Weiter wird untersucht, welche Rolle das Verhalten von Lehrkräften bei der Besprechung politischer Themen im Unterricht für die Genauigkeit und Richtung (d. h. Unter-/Überschätzung) der Selbsteinschätzung spielt.

Methoden. Die der Studie zugrundeliegenden Analysen basieren auf den Daten von 1.047 Schüler*innen der 7. ($n = 613$: $M_{\text{Alter}} = 12,47$; 47,4 % weiblich; 44,4 % Migrationshintergrund) und 10. Klassenstufe ($n = 434$: $M_{\text{Alter}} = 15,49$; 47,7 % weiblich; 35,9 % Migrationshintergrund). Die Daten wurden im Klassenkontext erfasst und stammen von insgesamt 69 Schulklassen aus NRW. In Anlehnung an die Originalstudie von Kruger und Dunning (1999) wurde der Zusammenhang zwischen dem selbsteingeschätzten und dem objektiv gemessenen politischen Wissen zunächst visuell überprüft und zwischen den Klassenstufen verglichen. Zusätzlich wurden zwei one-way und eine two-way ANOVA gerechnet, um zu überprüfen, ob sich die Differenz zwischen dem selbsteingeschätzten und dem objektiv gemessenen politischen Wissen zwischen den Leistungsquartilen und den beiden Klassenstufen unterscheidet. Schließlich wurde mit zwei Strukturgleichungsmodellen der Zusammenhang zwischen dem Verhalten von Lehrkräften und der Genauigkeit sowie der Richtung der Selbsteinschätzungen der Schüler*innen untersucht.

Ergebnisse. Die visuelle Überprüfung bestätigte das für den Dunning-Kruger Effekt typische Befundmuster in den Klassenstufen 7 und 10. Auch die beiden one-way ANOVAs zeigen, dass sich der Zusammenhang zwischen selbsteingeschätztem und objektiv gemessenem politischen Wissen statistisch signifikant zwischen den vier Leistungsquartilen in jeweils der 7. und 10. Klassenstufe unterschied. Weiter zeigt die visuelle Überprüfung, dass das Befundmuster in beiden Klassenstufen sehr ähnlich ausgeprägt war. Auch hier bestätigt eine two-way ANOVA den visuellen Eindruck, dass es statistisch signifikante Unterschiede zwischen den Leistungsquartilen, aber nicht zwischen den Klassenstufen gab. Abschließend demonstrieren die Strukturgleichungsmodelle, dass ein als motivierend wahrgenommener Unterricht mit einer genaueren Selbsteinschätzung, jedoch auch mit einer größeren Wahrscheinlichkeit der Überschätzung zusammenhing. Zudem hing ein als kognitiv aktivierend wahrgenommener Unterricht negativ mit der Genauigkeit von Selbsteinschätzungen zusammen.

Schlussfolgerung. Die Ergebnisse der Studie zeigen den Dunning-Kruger Effekt in Bezug auf das politische Wissen in sehr ähnlicher Form in der 7. und 10. Klassenstufe. Die Beständigkeit mit der sich der Effekt in zwei aus einer Entwicklungsperspektive so unterschiedlichen Kohorten zeigt, legen die Vermutung nahe, dass der Effekt tatsächlich ein statistisches Artefakt sein könnte. Zwar kann die Studie keine abschließende Bewertung der konkurrierenden Interpretationen des Dunning-Kruger Effekts leisten, jedoch verdeutlichen die Ergebnisse die Wichtigkeit einer kritischen Auseinandersetzung mit der klassischen metakognitiven Interpretation des Befunds. Außerdem verdeutlichen die Ergebnisse der Studie die Wichtigkeit von Lehrkräften und deren Unterrichtsgestaltung für die Selbsteinschätzungen von Schüler*innen im Kontext politischen Wissens.

Tabelle 1a*Veröffentlichte Beiträge der kumulativen Dissertation*

Beitrag	Autor*innen	Titel	Publikationsorgan
I	Alscher, P., Ludewig, U., & McElvany, N.	Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung	Zeitschrift für Erziehungswissenschaft
II	Alscher, P., Grecu, A. L., Ludewig, U., & McElvany, N.	Towards the Measurability of a Competence Model for School-based Civic Education	N. McElvany, A. L. Grecu, R. Lorenz, M. Becker, C. Dignath, H. Gaspard & F. Lauer mann (Hrsg.), 50 Jahre Schulentwicklung – Leitthemen der empirischen Bildungsforschung (S. 262–290). Beltz Juventa.
III	Ludewig, U., Alscher, P., & McElvany, N.	What Makes Domain Knowledge Difficult? Word Usage Frequency from SUBTLEX and dlexDB Explains Knowledge Item Difficulty	Behavior Research Methods
IV	Alscher, P., Ludewig, U., & McElvany, N.	Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators	Journal of Youth and Adolescence

Tabelle 1b

Beiträge zu weiterführenden Analysen

Beitrag	Autor*innen	Titel	Publikationsorgan
V	Alscher, P., Ludewig, U., Kleinkorres, R. & McElvany, N.	When Will They Know What They Don't Know? - Political knowledge and the Infamous Dunning-Kruger Effect	Englischsprachige Fachzeitschrift ^a

^aBeitrag V befindet sich aktuell in Revision.

Literaturverzeichnis I

- Aars, J. & Christensen, D. A. (2020). Education and political participation: the impact of educational environments. *Acta Politica*, 55(1), 86–102. <https://doi.org/10.1057/s41269-018-0101-5>
- Abs, H. J. & Hahn-Laudenberg, K. (Hrsg.). (2017). *Das politische Mindset von 14-Jährigen: Ergebnisse der International Civic and Citizenship Education Study 2016*. Waxmann.
- Achour, S. (2021). Gesellschaftliche Diversität: Herausforderungen und Ansätze. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 365–374). Wochenschau Verlag.
- Achour, S. & Meyer-Heidemann, C. (2020). Politische Bildung als Unterrichts- und Schulprinzip. In S. Achour, M. Busch, P. Massing & C. Meyer-Heidemann (Hrsg.), *Wörterbuch Politikunterricht* (S. 176–179). Wochenschau Verlag.
- Achour, S. & Wagner, S. (2019). *Wer hat, dem wird gegeben: Politische Bildung an Schulen: Bestandsaufnahme, Rückschlüsse und Handlungsempfehlungen* (Schriftenreihe des Netzwerk Bildung). Friedrich-Ebert-Stiftung.
- Ackerman, P. L. & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, 13(2), 57–78. <https://doi.org/10.1037/1076-898x.13.2.57>
- Aktionsrat Bildung. (2020). *Bildung zu demokratischer Kompetenz*. Waxmann. <https://doi.org/10.31244/9783830941811>
- Almond, G. A. & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.
- American Educational Research Association [AERA], American Psychological Association & National Council on Measurement in Education (2014). The standards for educational and psychological testing. American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf
- Anson, I. G. (2018). Partisanship, political knowledge, and the Dunning-Kruger effect. *Political Psychology*, 39(5), 1173–1192. <https://doi.org/10.1111/pops.12490>
- Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, 63(1), 18–25. <https://doi.org/10.1026/0033-3042/a000106>

- Asseburg, R. & Frey, A. (2013). Too hard, too easy, or just right? The relationship between effort or boredom and ability-difficulty fit. *Psychological Test and Assessment Modeling*, 55, 92–104.
- Autorengruppe Fachdidaktik (Hrsg.). (2011). *Politik und Bildung: Bd. 64. Konzepte der politischen Bildung: Eine Streitschrift*. Wochenschau Verlag.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Bayram-Özdemir, S., Stattin, H. & Özdemir, M. (2016). Youth's initiations of civic and political discussions in class: Do youth's perceptions of teachers' behaviors matter and why? *Journal of Youth and Adolescence*, 45(11), 2233–2245. <https://doi.org/10.1007/s10964-016-0525-z>.
- Behrmann, G. C., Grammes, T. & Reinhardt, S. (2004). politik: Kerncurriculum Sozialwissenschaften in der gymnasialen oberstufe. In H.-E. Tenorth (Hrsg.), *Kerncurriculum Oberstufe. 2. Biologie, Chemie, Physik, Geschichte, Politik. Expertisen - im Auftrag der KMK*. (S. 322–406). Beltz.
- Bermudez, A. (2015). Four tools for critical inquiry in history, social studies, and civic education. *Revista de estudios sociales*(52), 102–118. <https://doi.org/10.7440/res52.2015.07>
- Biesta, G. (2011). Warum „What works“ nicht funktioniert: Evidenzbasierte pädagogische Praxis und das Demokratiedefizit der Bildungsforschung. In J. Bellmann & T. Müller (Hrsg.), *Wissen, was wirkt: Kritik evidenzbasierter Pädagogik* (S. 95–121). Springer. https://doi.org/10.1007/978-3-531-93296-5_4
- Bittman, B. & Russell, W. (2016). Civic education in United States: A multiple regression of civic education scores from the National Assessment of Educational Progress. *Research in Social Sciences and Technology*, 1(2), 1–16. <https://doi.org/10.46303/ressat.01.02.1>
- Boese, V. A., Lundstedt, M., Morrison, K., Sato, Y. & Lindberg, S. I. (2022). State of the world 2021: autocratization changing its nature? *Democratization*, 29(6), 983–1013. <https://doi.org/10.1080/13510347.2022.2069751>
- Bohm, B. C. (2015). *Konzepte der Politik: Theoretische Grundlagen und unterrichtspraktische Herausforderung* [Dissertation]. Freie Universität Berlin, Berlin. https://refubium.fu-berlin.de/bitstream/handle/fub188/5303/Dissertation_Bohm.pdf;jsessionid=9E0A9CD5DE056A5F5C0699273BE202DB?sequence=1

- Bouffard, T., Vezeau, C., Roy, M. & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research*, 50(4), 221–229. <https://doi.org/10.1016/j.ijer.2011.08.003>
- Branigan, H. E. & Donaldson, D. I. (2020). Teachers matter for metacognition: Facilitating metacognition in the primary school through teacher-pupil interactions. *Thinking Skills and Creativity*, 38, 100718. <https://doi.org/10.1016/j.tsc.2020.100718>
- Brennan, J. (2016). *Against democracy*. Princeton University Press.
- Bromley, P. & Mäkinen, E. (2011). Diversity in Civic Education: Finland in Historical and Comparative Perspective. *Journal of International Cooperation in Education*, 14(2), 35–50.
- Brückner, S., Zlatkin-Troitschanskaia, O. & Pant, H. A. (2020). Standards für pädagogisches Testen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 217–248). Springer. https://doi.org/10.1007/978-3-662-61532-4_11
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J. & Böhl, A. (2011). The Word Frequency Effect. *Experimental Psychology*, 58(5), 412–424. <https://doi.org/10.1027/1618-3169/a000123>
- Brysbaert, M., Mandera, P. & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, 27(1), 45–50. <https://doi.org/10.1177/0963721417727521>
- Bundeswahlleiterin. (2023). *Wahlbeteiligung*. <https://www.bundeswahlleiter.de/service/glossar/w/wahlbeteiligung.html>
- Burden, B. C., Herd, P., Jones, B. M. & Moynihan, D. P. (2020). Education, early life, and political participation: New evidence from a sibling model. *Research & Politics*, 7(3). <https://doi.org/10.1177/2053168020958319>
- Campbell, D. E. (2019). What Social Scientists Have Learned About Civic Education: A Review of the Literature. *Peabody Journal of Education*, 94(1), 32–47. <https://doi.org/10.1080/0161956X.2019.1553601>
- Castillo, J. C., Miranda, D., Bonhomme, M., Cox, C. & Bascopé, M. (2015). Mitigating the political participation gap from the school: the roles of civic knowledge and classroom climate. *Journal of Youth Studies*, 18(1), 16–35. <https://doi.org/10.1080/13676261.2014.933199>
- Center for Civic Education. (1994). *National standards for civics and government*. Center for Civic Education.

- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. M.I.T. Press. <http://www.colinphilips.net/wp-content/uploads/2015/09/chomsky1965-ch1.pdf>
- Conrad, J., Lo, J. C. & Kisa, Z. (2022). Civic Mandates for the ‘Majority’: The Perception of Whiteness and Open Classroom Climate in Predicting Youth Civic Engagement. *The Journal of Social Studies Research*, 46(1), 7–17. <https://doi.org/10.1016/j.jssr.2021.11.005>
- Council of Europe. (2023). *Reference Framework of Competences for Democratic Culture (RFCDC)*. <https://www.coe.int/en/web/reference-framework-of-competences-for-democratic-culture>
- Crittenden, J. & Levine, P. (2018). Civic Education. In E. N. Zalta (Hrsg.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2023/entries/civic-education/>
- Dahl, R. A. (1989). *Democracy and Its Critics*. Yale University Press. <https://books.google.de/books?id=11RQngEACAAJ>
- Darling, N., Caldwell, L. L. & Smith, R. (2005). Participation in School-Based Extracurricular Activities and Adolescent Adjustment. *Journal of Leisure Research*, 37(1), 51–76. <https://doi.org/10.1080/00222216.2005.11950040>
- Del Chicca, L. & Maaß, J. (2014). Sparen mit Verstand – Möglichkeiten zur Vernetzung von Mathematik und politischer Bildung. In J. Maaß & H.-S. Siller (Hrsg.), *Neue Materialien für einen realitätsbezogenen Mathematikunterricht 2: ISTRON-Schriftenreihe* (S. 31–41). Springer. https://doi.org/10.1007/978-3-658-05003-0_3
- Detjen, J., Massing, P., Richter, D. & Weißeno, G. (2012). *Politikkompetenz – ein Modell*. Springer.
- Dewey, J. (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan. <https://books.google.de/books?id=OlqwHwAACAAJ>
- Dignath, C. & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231–264. <https://doi.org/10.1007/s11409-008-9029-x>
- Drippe, M. (2018). *Der Fachwortschatz in Schulbüchern für die 9. und 10. Klasse des Politikunterrichts an Gesamtschulen*. Cives! Scholl of Civic Education.
- Dunlosky, J. & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students’ learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>

- Dunning, D. (2011). Chapter five - The Dunning–Kruger Effect: On Being Ignorant of One's Own Ignorance. In J. M. Olson & M. P. Zanna (Hrsg.), *Advances in Experimental Social Psychology* (Bd. 44, S. 247–296). Academic Press. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Eckstein, K. & Noack, P. (2015). Classroom climate effects on adolescents' orientations toward political behaviors: a multilevel approach. In *Political Engagement of the Young in Europe* (S. 181–197). Routledge.
- Edelstein, W. & Fauser, P. (2001). „Demokratie lernen und leben“ Gutachten für ein Modellversuchsprogramm der BLK. Bonn. <http://www.blk-bonn.de/papers/heft96.pdf>
- Ekman, J. & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3), 283–300. <https://doi.org/10.2478/s13374-012-0024-1>
- Embretson, S. E. & Reise, S. P. *Item response theory*. Psychology Press. <https://doi.org/10.4324/9781410605269>
- Embse, N. von der, Jester, D., Roy, D. & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, 227, 483–493. <https://doi.org/10.1016/j.jad.2017.11.048>
- Fieldhouse, E., Tranmer, M. & Russel, A. (2007). Something about young people or something about elections? Electoral participation of young people in Europe: Evidence from a multilevel analysis of the European Social Survey. *European Journal of Political Research*, 46(6), 797–822. <https://doi.org/10.1111/j.1475-6765.2007.00713.x>
- Fischer, C. & Reinhardt, S. (2022). Bildung und Demokratie von Anfang an. Herausforderungen und Möglichkeiten. *Editorial zum ersten Sonderheft in itdb*, 4.
- Frank, S. (2005). *Demokratiebaustein: „Civic education“ - was ist das?* BLK 2005.
- Frech, S. & Richter, D. (2017). *Der Beutelsbacher Konsens: Bedeutung, Wirkung, Kontroversen*. Wochenschau Verlag.
- Freund, P. A. & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological bulletin*, 138(2), 296–321. <https://doi.org/10.1037/a0026556>
- Frey, A. & Hartig, J. (2013). Wann sollten computerbasierte Verfahren zur Messung von Kompetenzen anstelle von papier- und bleistift-basierten Verfahren eingesetzt werden? *Zeitschrift für Erziehungswissenschaft*, 16(1), 53–57. <https://doi.org/10.1007/s11618-013-0385-1>

- Gabriel, O. W. (2001). Politische Orientierungen im vereinigten Deutschland: Auf dem Weg zur „Civic Culture“? In H. Bertram & R. Kollmorgen (Hrsg.), *Die Transformation Ostdeutschlands* (S. 97–132). Springer.
- Gäde, J. C., Schermelleh-Engel, K. & Brandt, H. (2020). Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 615–659). Springer. https://doi.org/10.1007/978-3-662-61532-4_24
- Gagel, W. (2005). *Geschichte der politischen Bildung in der Bundesrepublik Deutschland 1945-1989/90*. Springer.
- Galston, W. A. (2007). Civic Knowledge, Civic Education, and Civic Engagement: A Summary of Recent Research. *International Journal of Public Administration*, 30(6-7), 623–642. <https://doi.org/10.1080/01900690701215888>
- Gerdes, J. (2022). Demokratiebildung. In U. Bauer, U. H. Bittlingmayer & A. Scherr (Hrsg.), *Handbuch Bildungs- und Erziehungssoziologie* (S. 1325–1350). Springer. https://doi.org/10.1007/978-3-658-30903-9_74
- Gignac, G. E. & Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data. *Intelligence*, 80, 101449. <https://doi.org/10.1016/j.intell.2020.101449>
- Gignac, G. E. & Zajenkowski, M. (2023). Still no Dunning-Kruger effect: A reply to Hiller. *Intelligence*, 97, 101733. <https://doi.org/10.1016/j.intell.2023.101733>
- Glaeser, E. L., Ponzetto, G. A. M. & Shleifer, A. (2007). Why does democracy need education? *Journal of Economic Growth*, 12(2), 77–99. <https://doi.org/10.1007/s10887-007-9015-1>
- Glavanovits, J., Gründl, J., Kritzinger, S. & Oberluggauer, P. (2019). Politische Partizipation. In J. Bacher, A. Grausgruber, M. Haller, F. Höllinger, D. Prandner & R. Verwiebe (Hrsg.), *Sozialstruktur und Wertewandel in Österreich: Trends 1986-2016* (S. 439–456). Springer. https://doi.org/10.1007/978-3-658-21081-6_18
- Gloe, M. & Oeftering, T. (2020). Didaktik der politischen Bildung: Ein Überblick über Ziele und Grundlagen inklusiver politischer Bildung. In D. Meyer, W. Hilpert & B. Lindmeier (Hrsg.), *Grundlagen und Praxis inklusiver politischer Bildung* (S. 87–92). Bundeszentrale für politische Bildung.
- Gogolin, I. (2020). Globale Bildungsakteure. In T. Hascher, T.-S. Idel & W. Helsper (Hrsg.), *Handbuch Schulforschung* (S. 1–15). Springer. https://doi.org/10.1007/978-3-658-24734-8_69-1

- Gökbudak, M., Hedtke, R. & Hagedorn, U. (2021). *4. Ranking Politische Bildung. Politische Bildung in der Sekundarstufe I und in der Berufsschule im Bundesländervergleich 2020*. Bielefeld: Universität, Fakultät für Soziologie.
- Goll, T. (2019). Die Rezeption des Modells der Politikkompetenz in der Politikdidaktik. In G. Weißeno (Hrsg.), *Politik lernen: Studien und theoretische Ansätze* (S. 25–39). Springer. https://doi.org/10.1007/978-3-658-27896-0_3
- Goll, T., Richter, D. & Weißeno, G. (2011). Politisches Wissen von Schüler/-innen mit und ohne Migrationshintergrund (POWIS): Ergebnisse einer Studie. In D. Lange (Hrsg.), *Schriftenreihe der DVPB, Wochenschau Wissenschaft, . Entgrenzungen: Gesellschaftlicher Wandel und politische Bildung* (S. 126–131). Wochenschau Verlag.
- Golz, H.-G. & Kost, A. (2021). Die Bundeszentrale und die Landeszentralen für politische Bildung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 194–203). Wochenschau Verlag.
- Goodboy, A. K. & Martin, M. M. (2020). Omega over alpha for reliability estimation of unidimensional communication measures. *Annals of the International Communication Association*, 44(4), 422–439. <https://doi.org/10.1080/23808985.2020.1846135>
- Gould, J., Hall Jamieson, K., Levine, P., McDonnell, T. & Smith, D. B. (2011). *Guardian of Democracy: The Civic Mission of Schools*. Pennsylvania. Leonore Annenberg Institute for Civics of the Annenberg Public Policy Center.
- GPJE. (2004). *Anforderungen an Nationale Bildungsstandards für den Fachunterricht in der Politischen Bildung an Schulen: Ein Entwurf*. Wochenschau Verlag. <http://gpje.de/wp-content/uploads/2017/01/Bildungsstandards-1.pdf>
- Graf, E., Goetz, T., Bieleke, M. & Murano, D. (2023). Feeling Politics at High School: Antecedents and Effects of Emotions in Civic Education. *Political Psychology*. Vorab-Onlinepublikation. <https://doi.org/10.1111/pops.12907>
- Grammes, T., Schluß, H. & Vogler, H.-J. (2006). *Staatsbürgerkunde in der DDR*. Springer. <https://link.springer.com/content/pdf/10.1007/978-3-531-90048-3.pdf>
- Grieger, M. (2022). Fächerübergreifendes Unterrichten im Fach Gesellschaftslehre. In *Selbstwirksamkeitserwartungen angehender und praktizierender Lehrkräfte zum Unterrichten von Gesellschaftslehre: Messung – Ausprägung – Prädiktoren* (S. 11–59). Springer. https://doi.org/10.1007/978-3-658-40155-9_2
- Grieger, M. & Oberle, M. (2020). Fächerübergreifendes Unterrichten im Fach Gesellschaftslehre: Herausforderungen und Ansätze für die Lehrerbildung. In A. Eis, G. Bade, A. Albrecht, U. Jakubczyk & B. Overwien (Hrsg.), *Jetzt erst recht: politische Bildung!*

- Bestandsaufnahme und bildungspolitische Forderungen* (S. 147–163). Wochenschau Verlag. <https://elibrary.utb.de/doi/epdf/10.46499/9783734409790>
- Grob, U. (2009). Die Entwicklung politischer Orientierung vom Jugend- ins Erwachsenenalter: - Ist die Jugend eine spezifisch sensible Phase in der politischen Sozialisation? In H. Fend, U. Grob & F. Berger (Hrsg.), *Lebensverläufe, Lebensbewältigung, Lebensglück: Ergebnisse der Life-Studie* (S. 305–326). Springer.
- Grobshäuser, N. (2022). Professionelle Kompetenz von Politiklehrkräften. In G. Weißeno & B. Ziegler (Hrsg.), *Handbuch Geschichts- und Politikdidaktik* (S. 283–298). Springer. https://doi.org/10.1007/978-3-658-29668-1_19
- Haarmann, M. P., Kenner, S. & Lange, D. (Hrsg.). (2020). *Demokratie, Demokratisierung und das Demokratische: Aufgaben und Zugänge der Politischen Bildung*. Springer.
- Hafener, B. (2018). Politische Bildung. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Handbuch Bildungsforschung* (S. 1111–1132). Springer. https://doi.org/10.1007/978-3-531-19981-8_49
- Hafener, B. (2021). Politische Bildung in der außerschulischen Jugendbildung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 176–185). Wochenschau Verlag.
- Hahn, C. (2010). Comparative civic education research: What we know and what we need to know. *Citizenship Teaching and Learning*, 6(1), 5–23. https://doi.org/10.1386/ctl.6.1.5_1
- Hahn, C. L. (2008). *The SAGE Handbook of Education for Citizenship and Democracy*. SAGE Publications Ltd. <https://doi.org/10.4135/9781849200486>
- Hahn-Laudenberg, K. (2022). Die gesellschaftliche Bedeutung der Politikdidaktik. In G. Weißeno & B. Ziegler (Hrsg.), *Handbuch Geschichts- und Politikdidaktik* (S. 77–92). Springer. https://doi.org/10.1007/978-3-658-29668-1_6
- Hahn-Laudenberg, K. & Abs, H. J. (2017). Politisches Wissen und Argumentieren. In H. J. Abs & K. Hahn-Laudenberg (Hrsg.), *Das politische Mindset von 14-Jährigen: Ergebnisse der International Civic and Citizenship Education Study 2016* (S. 77–111). Waxmann.
- Hartmann, H. & Thiery, P. (2022). *Global Findings BTI 2022*. Bertelsmann Stiftung. <https://doi.org/10.11586/2022032>
- Hartong, S. (2015). Global policy convergence through “distributed governance”? The emergence of “national” education standards in the US and Germany. *Journal of International and Comparative Social Policy*, 31(1), 10–33. <https://doi.org/10.1080/21699763.2014.977803>

- Hartong, S. (2018). *Standardbasierte Bildungsreformen in den USA*. Beltz Juventa. <https://www.beltz.de/fachmedien/erziehungswissenschaft/produkte/details/35832-standardbasierte-bildungsreformen-in-den-usa.html>
- Hedinger, F. (2018). Erkenntnisse aus der empirischen Politikwissenschaft für das Politikkompetenzmodell. In B. Ziegler & M. Waldis (Hrsg.), *Politische Bildung in der Demokratie: Interdisziplinäre Perspektiven* (S. 57–73). Springer. https://doi.org/10.1007/978-3-658-18933-4_5
- Hernández-Torrano, D. & Courtney, M. G. R. (2021). Modern international large-scale assessment in education: an integrative review and mapping of the literature. *Large-scale Assessments in Education*, 9(1), 17. <https://doi.org/10.1186/s40536-021-00109-1>
- Hiller, A. (2023). Comment on Gignac and Zajenkowski, “The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data”. *Intelligence*, 97, 101732. <https://doi.org/10.1016/j.intell.2023.101732>
- Hippe, T., Hedtke, R. & Hellmich, N. (2020). *Politische Bildung und Demokratie-Lernen in der Sekundarstufe. Curriculumanalyse für den Bereich Demokratiebildung für Kinder und Jugendliche in den Fächern der sozialwissenschaftlichen Bildung. Eine Expertise für das Deutsche Jugendinstitut*. Deutsches Jugendinstitut. https://www.dji.de/fileadmin/user_upload/dasdji/publikationen/Brosch%C3%BCren_2021_online/2020_Politische_Bildung_und_Demokratie_Lernen_in_der_Sekundarstufe.pdf
- Hofer, G., Mraulak, V., Grinschgl, S. & Neubauer, A. C. (2022). Less-intelligent and unaware? Accuracy and Dunning–Kruger effects for self-estimates of different aspects of intelligence. *Journal of Intelligence*, 10(1), 10. <https://doi.org/10.3390/jintelligence10010010>
- Höfling, V. & Moosbrugger, H. (2020). Standards für psychologisches Testen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 197–215). Springer. https://doi.org/10.1007/978-3-662-61532-4_10
- Holderberg, P. (2021). *Generation und politische Beteiligung in der Postdemokratie* [Dissertation]. Universität Hildesheim, Hildesheim.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. & Baird, J.-A. (2018). Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 62(3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>
- Jerome, L. (2012). Service learning and active citizenship education in England. *Education, Citizenship and Social Justice*, 7(1), 59–70. <https://doi.org/10.1177/1746197911432594>

- Jerrim, J. (2023). Has Peak PISA passed? An investigation of interest in International Large-Scale Assessments across countries and over time. *European Educational Research Journal*, 14749041231151793. <https://doi.org/10.1177/14749041231151793>
- Johansson, S. (2016). International large-scale assessments: what uses, what consequences? *Educational Research*, 58(2), 139–148. <https://doi.org/10.1080/00131881.2016.1165559>
- Juchler, I. (2010). Die Bedeutung von Basis- und Fachkonzepten für die kompetenzorientierte politische Bildung. In I. Juchler (Hrsg.), *Kompetenzen in der politischen Bildung* (S. 233–244). Wochenschau Verlag.
- Juhasz, B. J., Yap, M. J., Raoul, A. & Kaye, M. (2019). A Further Examination of Word Frequency and Age-of-Acquisition Effects in English Lexical Decision Task Performance: The Role of Frequency Trajectory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45, 82–96. <https://doi.org/10.1037/xlm0000564>
- Kenner, S. (2021). *Politische Bildung in Aktion: eine qualitative Studie zur Rekonstruktion von selbstbestimmten Bildungserfahrungen in politischen Jugendinitiativen*. Springer.
- Kenner, S. & Lange, D. (2020). Demokratiebildung. In S. Achour, M. Busch, P. Massing & C. Meyer-Heidemann (Hrsg.), *Wörterbuch Politikunterricht* (S. 48–51). Wochenschau Verlag.
- Klieme, E. (2022). Unterrichtsqualität. In M. Haring, C. Rohlfes & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 411–426). UTB. <https://elibrary.utb.de/doi/epdf/10.36198/9783838587967>
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Ri-quarts, K., Rost, J. & Tenorth, H.-E. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. BMBF.
- Klinger, T. (2022). Die Selbsteinschätzung von Sprachfähigkeiten: Eine Skala zur differenzier-ten Erfassung. In T. Klinger, I. Gogolin & B. Schnoor (Hrsg.), *Sprachentwicklung im Kontext von Mehrsprachigkeit: Hypothesen, Methoden, Forschungsperspektiven* (S. 79–112). Springer. https://doi.org/10.1007/978-3-658-36770-1_4
- Knekta, E. (2017). Are all Pupils Equally Motivated to do Their Best on all Tests? Differences in Reported Test-Taking Motivation within and between Tests with Different Stakes. *Scandinavian Journal of Educational Research*, 61(1), 95–111. <https://doi.org/10.1080/00313831.2015.1119723>

- Köller, O. (2016). Bildungsstandards. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Handbuch Bildungsforschung* (S. 1–24). Springer. https://doi.org/10.1007/978-3-531-20002-6_26-1
- Köller, O. (2020). Bildungsstandards. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Springer Reference Sozialwissenschaften. Handbuch Bildungsforschung* (Living reference work, continuously updated edition). Springer. https://doi.org/10.1007/978-3-531-20002-6_26-1
- Kostelka, F. (2017). Does Democratic Consolidation Lead to a Decline in Voter Turnout? Global Evidence Since 1939. *The American Political Science Review*, *111*(4), 653–667. <https://doi.org/10.1017/S0003055417000259>
- Krösche, H. (2020). Die Sicht der Kinder ins Zentrum rücken. Zum Stellenwert des politischen Lernens in der österreichischen Primarstufe. In M. P. Haarmann, S. Kenner & D. Lange (Hrsg.), *Demokratie, Demokratisierung und das Demokratische: Aufgaben und Zugänge der Politischen Bildung* (S. 233–249). Springer. https://doi.org/10.1007/978-3-658-29556-1_16
- Krueger, J. & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, *82*(2), 180.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. <https://doi.org/10.1037//0022-3514.77.6.1121>
- Kuger, S. & Gille, M. (2020). Entwicklung des zivilgesellschaftlichen Engagements im Jugend- und jungen Erwachsenenalter. *Zeitschrift für Erziehungswissenschaft*, *23*(6), 1103–1123. <https://doi.org/10.1007/s11618-020-00978-2>
- Kuhn, H.-J. (2014). Anspruch, Wirklichkeit und Perspektiven der Gesamtstrategie der KMK zum Bildungsmonitoring. *DDS–Die Deutsche Schule*, *106*(4), 414–426.
- Kuhn, H.-W., Massing, P. & Skuhr, W. (1993). Politische Bildung im Nationalsozialismus. In H.-W. Kuhn, P. Massing & W. Skuhr (Hrsg.), *Politische Bildung in Deutschland: Entwicklung — Stand — Perspektiven* (S. 81–107). Springer. https://doi.org/10.1007/978-3-663-05736-9_3
- Kühne, R. (2013). Konzeptspezifikation und Messung. In W. Möhring & D. Schlütz (Hrsg.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* (S. 23–40). Springer. https://doi.org/10.1007/978-3-531-18776-1_2

- Kultusministerkonferenz. (2018). *Demokratie als Ziel, Gegenstand und Praxis historisch-politischer Bildung und Erziehung in der Schule*. KMK. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2009/2009_03_06-Staerkung_Demokratieerziehung.pdf
- Kunter, M. & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Waxmann Verlag.
- Kunter, M., Klusmann, U. & Baumert, J. (2009). Professionelle Kompetenz von Mathematiklehrkräften: Das COACTIV-Modell. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 153–165). Beltz.
- Landwehr, B. (Hrsg.). (2017). *Partizipation, Wissen und Motivation im Politikunterricht*. Springer. <https://doi.org/10.1007/978-3-658-16507-9>
- Langton, K. P. & Jennings, M. K. (1968). Political Socialization and the High School Civics Curriculum in the United States. *The American Political Science Review*, 62(3), 852–867. <https://doi.org/10.2307/1953435>
- Lau, J. Y. F. (2019). Reflections on the Umbrella Movement: Implications for civic education and critical thinking. *Educational Philosophy and Theory*, 51(2), 163–174. <https://doi.org/10.1080/00131857.2017.1310014>
- Le, K. & Nguyen, M. (2021). Education and political engagement. *International Journal of Educational Development*, 85, 102441. <https://doi.org/10.1016/j.ijedudev.2021.102441>
- Lenzi, M., Vieno, A., Sharkey, J., Mayworm, A., Scacchi, L., Pastore, M. & Santinello, M. (2014). How school can teach civic engagement besides civic education: The role of democratic school climate. *American journal of community psychology*, 54(3-4), 251–261.
- Ling, G., Attali, Y., Finn, B. & Stone, E. A. (2017). Is a Computerized Adaptive Test More Motivating Than a Fixed-Item Test? *Applied Psychological Measurement*, 41(7), 495–511. <https://doi.org/10.1177/0146621617707556>
- Liu, J. & Dorans, N. J. (2016). Fairness in Score Interpretation. In N. J. Dorans & L. Cook (Hrsg.), *Fairness in Educational Assessment and Measurement* (S. 77–96). Routledge.
- Maag Merki, K. (2016). Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen

- Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 151–181). Springer. https://doi.org/10.1007/978-3-531-18942-0_6
- Magnus, J. R. & Peresetsky, A. A. (2022). A Statistical Explanation of the Dunning-Kruger Effect. *Frontiers in Psychology*, *13*, 840180. <https://doi.org/10.3389/fpsyg.2022.840180>
- Malak-Minkiewicz, B. & Torney-Purta, J. (2021). *Influences of the IEA civic and citizenship education studies: Practice, policy, and research across countries and regions*. Springer.
- Malkewitz, C. P., Schwall, P., Meesters, C. & Hardt, J. (2023). Estimating reliability: A comparison of Cronbach's α , McDonald's ω t and the greatest lower bound. *Social Sciences & Humanities Open*, *7*(1), 100368. <https://doi.org/10.1016/j.ssaho.2022.100368>
- Manganelli, S., Lucidi, F. & Alivernini, F. (2015). Italian adolescents' civic engagement and open classroom climate: The mediating role of self-efficacy. *Journal of Applied Developmental Psychology*, *41*, 8–18. <https://doi.org/10.1016/j.appdev.2015.07.001>
- Martens, A. M. & Gainous, J. (2013). Civic education and democratic capacity: How do teachers teach and what works? *Social science quarterly*, *94*(4), 956–976. <https://doi.org/10.1111/j.1540-6237.2012.00864.x>
- Massing, P. (2022). Kompetenzmodelle in der Politikdidaktik. In G. Weißeno & B. Ziegler (Hrsg.), *Handbuch Geschichts- und Politikdidaktik* (S. 17–32). Springer. https://doi.org/10.1007/978-3-658-29668-1_2
- Maurissen, L. (2020). Political efficacy and interest as mediators of expected political participation among Belgian adolescents. *Applied Developmental Science*, *24*(4), 339–353. <https://doi.org/10.1080/10888691.2018.1507744>
- McElvany, N. (2022). Cross-national achievement surveys and educational monitoring in Germany. In European Commission, Joint Research Centre, L. Volante, S. Schnepf, D. Klinger (Hrsg.), *Cross-national achievement surveys for monitoring educational outcomes : policies, practices and political reforms within the European Union* (S. 36-50). Publications Office of the European Union. <https://data.europa.eu/doi/10.2760/406165>
- McIntosh, R. D., Moore, A. B., Liu, Y. & Della Sala, S. (2022). Skill and self-knowledge: empirical refutation of the dual-burden account of the Dunning–Kruger effect. *Royal Society Open Science*, *9*(12), 191727.
- Memon, M. A., Joughin, G. R. & Memon, B. (2010). Oral assessment and postgraduate medical examinations: establishing conditions for validity, reliability and fairness. *Advances*

- in Health Sciences Education*, 15(2), 277–289. <https://doi.org/10.1007/s10459-008-9111-9>
- Messick, S. (1984). The Psychology of Educational Measurement. *Journal of Educational Measurement*, 21(2), 215–237.
- Milner, H. (2002). *Civic Literacy: How Informed Citizens Make Democracy Work*. University Press of New England. <https://books.google.de/books?id=zy2GAAAAMAAJ>
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2022). *Das Schulwesen in NRW aus quantitativer Sicht 2021/22: Statistische Übersicht Nr. 417*. MKW. https://www.schulministerium.nrw/system/files/media/document/file/quantita_2021.pdf
- Mirra, N. & Garcia, A. (2020). “I hesitate but I do have hope”: Youth speculative civic literacies for troubled times. *Harvard Educational Review*, 90(2), 295–321.
- Moosbrugger, H. & Brandt, H. (2020). Itemkonstruktion und Antwortverhalten. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 67–89). Springer. https://doi.org/10.1007/978-3-662-61532-4_4
- Moosbrugger, H., Gäde, J. C., Schermelleh-Engel, K. & Rauch, W. (2020). Klassische Testtheorie (KTT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 275–304). Springer. https://doi.org/10.1007/978-3-662-61532-4_13
- Moosbrugger, H. & Kelava, A. (2020). Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien“). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 13–38). Springer. https://doi.org/10.1007/978-3-662-61532-4_2
- Mulder, M., Weigel, T. & Collins, K. (2007). The concept of competence in the development of vocational education and training in selected EU member states: a critical analysis. *Journal of Vocational Education & Training*, 59(1), 67–88. <https://doi.org/10.1080/13636820601145630>
- Munniksma, A., Ziemes, J. & Jugert, P. (2021). Ethnic Diversity and Students’ Social Adjustment in Dutch Classrooms. *Journal of Youth and Adolescence*, 51, 141–155. <https://doi.org/10.1007/s10964-021-01507-y>
- Nagengast, B. & Rose, N. (2020). Quantitative Bildungsforschung und Assessments. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Springer Reference Sozialwissenschaften. Handbuch Bildungsforschung* (S. 1–20). Springer. https://doi.org/10.1007/978-3-531-20002-6_28-1

- National Assessment Governing Board. (2018). *Civics Framework: for the 2018 National Assessment of Educational Progress*. Washington, DC. National Assessment Governing Board, U.S. Department of Education. <https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/civics/2018-civics-framework.pdf>
- National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.
- Niemi, R. G. & Junn, J. (1998). *Civic education: What makes students learn*. Yale University Press. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=52882>
- Oberle, M., Ivens, S. & Leunig, J. (2018). Grenzenlose Toleranz? Lehrervorstellungen zum Beutelsbacher Konsens und dem Umgang mit Extremismus im Unterricht. In L. Möllers & S. Manzel (Hrsg.), *Populismus und politische Bildung* (S. 53–61). Wochenschau Verlag.
- Oberle, M. & Schwanholz, J. (2023). *Demokratie ohne Wahlbeteiligung? Potenziale politischer Bildung an Schulen besser nutzen!* Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/bueros/nrw/19927.pdf>
- Oberle, M. & Weißeno, G. (Hrsg.). (2017). *Politikwissenschaft und Politikdidaktik: Theorie und Empirie*. Springer.
- Oberle, M. & Wenzel, N. (2019). Politisches Vertrauen und Effektivitätsgefühl von Schüler/innen - Einflussfaktoren und Relevanz für politische Partizipation. In M. Lotz & K. Pohl (Hrsg.), *Schriftenreihe der GPJE. Gesellschaft im Wandel: Neue Aufgaben für die politische Bildung und ihre Didaktik* (S. 176–186). Wochenschau Verlag.
- Ohtani, K. & Hisasaka, T. (2018). Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, 13(2), 179–212. <https://doi.org/10.1007/s11409-018-9183-8>
- Partnership for 21st Century Skills. (2009). *P21 framework definitions*. <https://files.eric.ed.gov/fulltext/ED519462.pdf>
- Penfield, R. D. (2016). Fairness in Test Scoring. In N. J. Dorans & L. Cook (Hrsg.), *Fairness in Educational Assessment and Measurement* (S. 55–75). Routledge. <https://doi.org/10.4324/9781315774527-11>
- Perry, J., Lundie, D. & Golder, G. (2019). Metacognition in schools: what does the literature suggest about the effectiveness of teaching metacognition in schools? *Educational Review*, 71(4), 483–500. <https://doi.org/10.1080/00131911.2018.1441127>

- Pohl, K. (2021). Politische Bildung als Unterrichtsfach in der Schule. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 143–151). Wochenschau Verlag.
- Praetorius, A.-K. & Gräsel, C. (2021a). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>
- Praetorius, A.-K. & Gräsel, C. (Hrsg.). (2021b). Unterrichtsqualität – fachspezifisch oder generisch? Ein Update zu Perspektiven aus der Fachdidaktik und der allgemeinen Lehr-Lernforschung [Sonderheft]. *Unterrichtswissenschaft*, 49(2). Springer.
- 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., Lauermann, F., Klassen, R. M., Dickhäuse, O., Janke, S. & Dresel, M. (2017). Longitudinal relations between teaching-related motivations and student-reported teaching quality. *Teaching and Teacher Education*, 65, 241–254. <https://doi.org/10.1016/j.tate.2017.03.023>
- Prenzel, M. (2018). Rückwirkungen von Large Scale Assessments auf das Bildungssystem. *RdJB Recht der Jugend und des Bildungswesens*, 66(1), 27–39.
- Rapeli, L. (2023). Illusion of knowledge: is the Dunning-Kruger effect in political sophistication more widespread than before? *Journal of Elections, Public Opinion and Parties*, 1–11. <https://doi.org/10.1080/17457289.2023.2214734>
- Rauch, D. & Hartig, J. (2020). Interpretation von Testwerten in der Item-Response-Theorie (IRT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 411–424). Springer. https://doi.org/10.1007/978-3-662-61532-4_17
- Reichhart, B. (2018). *Lehrerprofessionalität im Bereich der politischen Bildung*. Springer.
- Reinders, H. & Post, I. (2022). Testverfahren. In H. Reinders, D. Bergs-Winkels, A. Prochnow & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 175–193). Springer. https://doi.org/10.1007/978-3-658-27277-7_11
- Reusser, K. & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49(2), 189–202. <https://doi.org/10.1007/s42010-021-00117-8>
- Rios, J. (2021). Improving Test-Taking Effort in Low-Stakes Group-Based Educational Testing: A Meta-Analysis of Interventions. *Applied Measurement in Education*, 34(2), 85–106. <https://doi.org/10.1080/08957347.2021.1890741>

- Rust, J. & Golombok, S. (2014). *Modern psychometrics: The science of psychological assessment*. Routledge.
- Sanborn, H. & Thyne, C. L. (2014). Learning Democracy: Education and the Fall of Authoritarian Regimes. *British Journal of Political Science*, 44(4), 773–797. <https://doi.org/10.1017/S0007123413000082>
- Sander, W. (2020). Fächerintegration in den Gesellschaftswissenschaften: Neue Ansätze und theoretische Grundlagen. In A. Eis, G. Bade, A. Albrecht, U. Jakubczyk & B. Overwien (Hrsg.), *Jetzt erst recht: politische Bildung! Bestandsaufnahme und bildungspolitische Forderungen* (S. 98–114). Wochenschau Verlag.
- Sander, W. (2021a). Geschichte der politischen Bildung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 13–28). Wochenschau Verlag.
- Sander, W. (2021b). Kompetenzorientierung als Forschungs- und Konfliktfeld der Didaktik der politischen Bildung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 122–132). Wochenschau Verlag.
- Scheunpflug, A. (2021). Global Citizenship Education: Weltgesellschaft als Herausforderung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 410–418). Wochenschau Verlag.
- Schlösser, T., Dunning, D., Johnson, K. & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the “signal extraction” counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, 39, 85–100. <https://doi.org/10.1016/j.joep.2013.07.004>
- Schneider, H. & Gerold, M. (2018). *Demokratiebildung an Schulen – Analyse lehrerbezogener Einflussgrößen*. Bertelsmann Stiftung.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B. & Agrusti, G. (2016). *IEA International Civic and Citizenship Education Study 2016: Assessment Framework*. Springer.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G. & Friedman, T. (2018). *Becoming Citizens in a Changing World: IEA International Civic and Citizenship Education Study 2016: International Report*. Springer. <https://doi.org/10.1007/978-3-319-73963-2>
- Schulz, W., Fraillon, J., Losito, B., Agrusti, G., Ainley, J., Damiani, V. & Friedman, T. (2022). *International Civic and Citizenship Education Study 2022: Assessment Framework*. Springer.
- Sheldrake, R., Mujtaba, T. & Reiss, M. J. (2022). Implications of under-confidence and over-confidence in mathematics at secondary school. *International Journal of Educational Research*, 116, 102085. <https://doi.org/10.1016/j.ijer.2022.102085>

- Sondheimer, R. M. & Green, D. P. (2010). Using Experiments to Estimate the Effects of Education on Voter Turnout. *American Journal of Political Science*, 54(1), 174–189. <http://www.jstor.org/stable/20647978>
- Spoden, C., Fleischer, J. & Frey, A. (2020). Person misfit, test anxiety, and test-taking motivation in a large-scale mathematics proficiency test for self-evaluation. *Studies in Educational Evaluation*, 67, 100910. <https://doi.org/10.1016/j.stueduc.2020.100910>
- Stanovich, K. E. & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of educational psychology*, 85, 211–229. <https://api.semanticscholar.org/CorpusID:199185040>
- Stevenson, J. (1995). The Political Colonisation of the Cognitive Construction of Competence. *The Vocational Aspect of Education*, 47(4), 353–364. <https://doi.org/10.1080/0305787950470402>
- Tampio, N. (2017). Democracy and national education standards. *The Journal of Politics*, 79(1), 33–44. <https://doi.org/10.1086/687206>
- Torney-Purta, J., Schwille, J. & Amadeo, J.-A. (Hrsg.). (1999a). *Civic education across countries: Twenty-four national case studies from the IEA civic education project*. Eburon Publishers. <https://files.eric.ed.gov/fulltext/ED431705.pdf>
- Torney-Purta, J., Schwille, J. & Amadeo, J.-A. (1999b). Mapping the Distinctive and Common Features of Civic Education in Twenty-Four Countries. In J. Torney-Purta, J. Schwille & J.-A. Amadeo (Hrsg.), *Civic education across countries: Twenty-four national case studies from the IEA civic education project* (S. 11–36). Eburon Publishers.
- Tsatsanis, E., Andreadis, I. & Teperoglou, E. (2018). Populism from Below: Socio-economic and Ideological Correlates of Mass Attitudes in Greece. *South European Society and Politics*, 23(4), 429–450. <https://doi.org/10.1080/13608746.2018.1510635>
- van der Stel, M. & Veenman, M. V. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, 20(3), 220–224. <https://doi.org/10.1016/j.lindif.2009.11.005>
- van Helvoort, J. (2019). Four spaces of civic literacy education: A literature review. In S. Kurbanoglu, S. Špiranec, Y. Ünal, J. Boustany, M. L. Huotari, E. Grassian, D. Mizrachi & L. Roy (Hrsg.), *Information Literacy in Everyday Life* (S. 94–102). Springer.
- van Prooijen, J.-W. (2021). Overconfidence in radical politics. In J. P. Forgas, W. D. Crano & K. Fiedler (Hrsg.), *The psychology of populism* (S. 143–157). Taylor & Francis.
- Velez, G. M. & Knowles, R. T. (2022). Trust, civic self-efficacy, and acceptance of corruption among Colombian adolescents: shifting attitudes between 2009-2016. *Compare: A*

- Journal of Comparative and International Education*, 52(8), 1205–1221.
<https://doi.org/10.1080/03057925.2020.1854084>
- Vonken, M. (2017). Competence, Qualification and Action Theory. In M. Mulder (Hrsg.), *Competence-based Vocational and Professional Education: Bridging the Worlds of Work and Education* (S. 67–82). Springer. https://doi.org/10.1007/978-3-319-41713-4_3
- Vuorre, M. & Metcalfe, J. (2022). Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacognition and Learning*, 17(2), 269–291. <https://doi.org/10.1007/s11409-020-09257-1>
- Waldis, M. (2020). „Civic media literacy“, „digital citizenship“ und Politische Kompetenz – Annäherungen an ein theoretisches Rahmenmodell für die digitale Politische Bildung. In U. Binder & J. Drerup (Hrsg.), *Demokratieerziehung und die Bildung digitaler Öffentlichkeit* (S. 55–70). Springer. https://doi.org/10.1007/978-3-658-28169-4_4
- Weinberg, J. (2022). Civic education as an antidote to inequalities in political participation? New evidence from English secondary education. *British Politics*, 17(2), 185–209. <https://doi.org/10.1057/s41293-021-00186-4>
- Weißeno, G. (2007). Kompetenzmodell. In G. Weißeno, K.-P. Hufer, H.-W. Kuhn, P. Massing & D. Richter (Hrsg.), *Wörterbuch Politische Bildung* (S. 175–181). Wochenschau Verlag.
- Weißeno, G. (2018). Politikdidaktik als Wissenschaftsdisziplin. In G. Weißeno, R. Nickolaus, M. Oberle & S. Seeber (Hrsg.), *Gesellschaftswissenschaftliche Fachdidaktiken: Theorien, empirische Fundierungen und Perspektiven* (S. 21–37). Springer. https://doi.org/10.1007/978-3-658-18892-4_2
- Weißeno, G. (2022). Bildungsstandards in der Politischen Bildung. In V. Reinhardt & D. Lange (Hrsg.), *Basiswissen Politische Bildung: Forschung, Planung und Methoden Politischer Bildung* (S. 105–126). Schneider.
- Weißeno, G., Detjen, J., Juchler, I., Massing, P. & Richter, D. (2010). *Konzepte der Politik - ein Kompetenzmodell. Schriftenreihe: Bd. 1016*. Bundeszentrale für politische Bildung. https://www.pedocs.de/volltexte/2016/12009/pdf/Weisseno_et_al_2010_Konzepte_der_Politik_.pdf
- Weißeno, G. & Grobshäuser, N. (2021). Unterrichtsqualität des Politikunterrichts in der gymnasialen Oberstufe. In C. Deichmann & M. Partetzke (Hrsg.), *Demokratie im Stresstest: Reaktionen von Politikdidaktik und politischer Bildung* (S. 249–264). Springer. https://doi.org/10.1007/978-3-658-33077-4_14

- Weißeno, G. & Landwehr, B. (2015). Knowledge about the European Union in Political Education: What are the Effects of Motivational Predispositions and Cognitive Activation? *McGill Journal of Education / Revue des sciences de l'éducation de McGill*, 50(2-3), 413–432.
- Weißeno, G., Weschenfelder, E. & Oberle, M. (2013). Empirische Ergebnisse zur professionellen Kompetenz von Politiklehrer/-innen. In K.-P. Hufer & D. Richter (Hrsg.), *Politische Bildung als Profession–Verständnisse und Forschungen* (S. 187–202). Bundeszentrale für politische Bildung.
- Welling, J., Gnams, T. & Carstensen, C. H. (2023). Identifying Disengaged Responding in Multiple-Choice Items: Extending a Latent Class Item Response Model With Novel Process Data Indicators. *Educational and Psychological Measurement*, 00131644231169211. <https://doi.org/10.1177/00131644231169211>
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Winter, T. (2021). Politische Partizipation in Parteien. In F. Bätge, K. Effing, K. Möltgen-Sicking & T. Winter (Hrsg.), *Politische Partizipation* (S. 69–82). Springer. https://doi.org/10.1007/978-3-658-33985-2_4
- Winther, E. (2020). Kompetenzerfassung und -entwicklung in der Bildungsforschung. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Springer Reference Sozialwissenschaften. Handbuch Bildungsforschung* (S. 1–16). Springer. https://doi.org/10.1007/978-3-531-20002-6_46-1
- Wise, S. L. (2017). Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications. *Educational Measurement: Issues and Practice*, 36(4), 52–61. <https://doi.org/10.1111/emip.12165>
- Yang, J. & Hoskins, B. (2020). Does university have an effect on young people's active citizenship in England? *Higher Education*, 80(5), 839–856. <https://doi.org/10.1007/s10734-020-00518-1>
- Yang, S. C. & Chung, T.-Y. (2009). Experimental study of teaching critical thinking in civic education in Taiwanese junior high school. *British Journal of Educational Psychology*, 79(1), 29–55.
- Zell, E. & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science*, 9(2), 111–125. <https://doi.org/10.1177/1745691613518075>

Zhou, X. & Jenkins, R. (2020). Dunning–Kruger effects in face perception. *Cognition*, 203, 104345. <https://doi.org/10.1016/j.cognition.2020.104345>

Ziegler, B. (2018). Politische Bildung im fächerübergreifenden Unterricht: Gefestigte Perspektiven und fachliche Konzepte als Grundlage. In S. Manzel & M. Oberle (Hrsg.), *Kompetenzorientierung* (S. 35–46). Springer .

2 Beiträge der kumulativen Dissertation

2.1 Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung

Alscher, P., Ludewig, U. & McElvany, N. (2022). Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung. *Zeitschrift für Erziehungswissenschaft*, 25(5), 1221–1241. <https://doi.org/10.1007/s11618-022-01085-0>

Dieser Artikel stellt keine Kopie der Veröffentlichung dar und entspricht möglicherweise nicht exakt der endgültigen und maßgeblichen Version des in der Zeitschrift veröffentlichten Artikels.

Zusammenfassung

Das Funktionieren einer Demokratie beruht nicht allein auf Institutionen und Gesetzen, sondern auch auf dem Engagement der Bürger*innen und der diesem Engagement zugrundeliegenden Civic Literacy. Die theoretische Kompetenzforschung in der schulischen politischen Bildung unterscheidet zwischen politischem Wissen sowie Handlungsdispositionen in den Bereichen Motivation, Einstellung und Volition. Das Zusammenwirken dieser Bereiche stellt eine Herausforderung für die empirische Kompetenzforschung dar. Nur durch das gemeinsame Messen und Interpretieren sowie das separate Auswerten dieser verschiedenen Bereiche wird eine differenzierte Betrachtung von Civic Literacy möglich. Vor diesem Hintergrund wird aufbauend auf Arbeiten aus der Politikdidaktik, Erziehungswissenschaft und Psychologie im folgenden Beitrag ein theoriebasiertes Kompetenzmodell, welches im Kern einen politischen Wissenstest und darüber hinaus Handlungsdispositionen in den Bereichen Motivation, Einstellung und Volition enthält, vorgestellt. Das Kompetenzmodell soll eine theoriebasierte, systematische Beforschung und Evaluierung von Civic Literacy bei Schüler*innen für Forschung und Praxis ermöglichen.

Schlüsselwörter: Civic Literacy · Kompetenzentwicklung · Politische Bildung · Schule

Abstract

Democratic institutions and laws are rarely sufficient to ensure the survival of democracy. Instead, democracies can only function well if embedded in a culture of democracy. Therefore, the quality and persistence of a democracy depends on the engagement and the underlying civic literacy of its citizenry. Civic education research differentiates between political knowledge and motivational, attitudinal, and volitional dispositions. By measuring all these components and carefully interpreting their interactions we obtain a more sophisticated and differentiated understanding of civic literacy in youth. Accordingly, this paper conceptualizes a theoretically and empirically justified competence model, which addresses children's political knowledge as well as their motivational, attitudinal, and volitional dispositions. In the future, this competence model will allow for theoretically sound and systematic research on civic literacy in youth and provide valuable insights for research and practice in the field.

Keywords: Civic Literacy · Competence development · Political education · School

Einleitung

Nicht zuletzt der sogenannte Sturm auf das US-amerikanische Kapitol im Januar 2021 hat erneut gezeigt, dass die Demokratie keine Selbstverständlichkeit ist und vielerorts durch demokratiefeindliche Bewegungen aus unterschiedlichen Richtungen bedroht wird. Dass die Geschehnisse in den USA hierfür kein Einzelfall sind, zeigt ein Vorfall aus dem August 2020 in Berlin, als Rechtsradikale mit schwarz-weiß-roten Reichsflaggen die Treppe des Bundestages stürmten, versuchten in den Bundestag einzudringen und dabei Polizistinnen und Polizisten bedrohten und teilweise gewaltvoll angriffen. Gleichzeitig ermöglichen digitale Medien den niedrigschwelligen, häufig anonymen Austausch sowie die Vernetzung politisch Gleichgesinnter in verschiedenen (sozialen) Bewegungen, z. B. „Fridays for Future“, „Black Lives Matter“, „Pulse of Europe“, aber auch „PEGIDA“ oder „Querdenken“. Dies stellt Gesellschaften und mit ihr die politischen Entscheidungsträgerinnen und -träger sowie die Wissenschaft beziehungsweise die Politikdidaktik vor die Herausforderung, eine (schulische) politische Bildung zu ermöglichen und zu fördern, die junge Menschen für politische und gesellschaftliche Partizipation begeistert, dabei Meinungsvielfalt und -freiheit erlaubt und gleichzeitig demokratische Strukturen und Institutionen schützt.

Das Fehlen beziehungsweise der Zerfall demokratischer Strukturen und Institutionen führt zum Exitus einer Demokratie; demokratische Strukturen und Institutionen alleine reichen aber auch nicht aus, um das Überleben einer Demokratie zu sichern (siehe Almond und Verba, 1963; Putnam, 2000; Tocqueville, 1838). Vielmehr beruht das Funktionieren einer Demokratie auch auf dem Engagement der Bürgerinnen und Bürger und der diesem Engagement zugrundeliegenden politischen und zivilgesellschaftlichen Kompetenz, welche im weiteren Verlauf dieses Beitrags als Civic Literacy beschrieben wird. Trotz der Vielschichtigkeit des Civic Literacy Begriffes beschreiben vor allem die folgenden zwei Merkmale die eigentliche Bedeutung des Begriffes: Civic Literacy beinhaltet (1) das Wissen und die Fähigkeiten, informierte Entscheidungen hinsichtlich gesellschaftlich relevanter Fragestellungen zu treffen, und (2) die für die Umsetzung der getroffenen Entscheidungen prinzipiell notwendigen Handlungsdispositionen (d. h. Motivationen, Einstellungen und Volitionen) zur politischen und zivilgesellschaftlichen Partizipation (z. B. Kultusministerkonferenz, 2018; Milner, 2002; Partnership for 21st Century Skills, 2009).

Das politische Wissen sowie die für Partizipation notwendigen Handlungsdispositionen müssen, auch in einer gefestigten Demokratie, von jeder Generation neu erlernt beziehungsweise verhandelt werden (siehe Edelstein und Fauser, 2001). Welches Wissen und welche

Handlungsdispositionen die heranwachsende Generation von Bürgerinnen und Bürgern in Bezug auf die Demokratie und die Gesellschaft entwickelt, ist daher von besonderer Relevanz. Wichtig ist, dass die Vermittlung dieser Fähigkeiten kohärent, chancengleich und auf Basis gemeinsamer Werte und Normen erfolgt.

Aktuelle Befunde zeigten jedoch, dass schon im Jugendalter das Phänomen einer politischen Ungleichheit entlang der sozialen Ungleichheit zu beobachten ist (Achour und Wagner, 2019). Herkunftsbedingte und systematische Ungleichheiten im politischen Wissen (und in politischen Einstellungen) „beschreiben [...] einen Aspekt der Reproduktion sozialer und politischer Ungleichheiten und haben nicht nur eine sozialwissenschaftsdidaktische, sondern auch eine politische Dimension“ (Hahn-Laudenberg, 2020, S. 53). Ergebnisse der International Civic and Citizenship Education Study (ICCS) 2016 verdeutlichten, dass das politische Wissen von 14-jährigen Schülerinnen und Schülern in Nordrhein-Westfalen in relativ starker Abhängigkeit vom kulturellen Kapital des Elternhauses steht. Unter den 16 europäischen Vergleichsländern ist diese Abhängigkeit nur in Bulgarien noch stärker (Abs und Hahn-Laudenberg, 2017). Allerdings muss die Bedeutung dieser Ergebnisse für den gesamtdeutschen Kontext, aufgrund der bildungsföderalismusbedingten Unterschiede zwischen den Bundesländern, sorgfältig abgewogen werden. Nichtsdestotrotz deuten die Ergebnisse aus dem bevölkerungsreichsten Bundesland auf den Bedarf einer systematischen Erfassung hinsichtlich des Erwerbs einer Civic Literacy im Rahmen der schulischen politischen Bildung hin, um Ungleichheiten und weitere Defizite erkennen, erklären und perspektivisch ausgleichen zu können.

Der vorliegende Beitrag konzeptualisiert ein Rahmenmodell, welches die systematische Beforschung und Evaluierung von Civic Literacy bei Schülerinnen und Schülern ermöglichen soll. Zudem wird ein neuentwickeltes Testinstrument für die Erfassung des politischen Wissens vorgestellt. Anders als bereits existierende Tests zum politischen Wissen, berücksichtigt dieser Besonderheiten des politischen Systems der Bundesrepublik Deutschland, folgt einer systematischen Konzeptualisierung von drei Inhaltsbereichen und drei intellektuellen Fähigkeiten und ermöglicht zudem die Erfassung des politischen Wissens im Längsschnitt.

Politische Bildung in und durch Schulen

Eine Schlüsselrolle bei der chancengleichen und kohärenten Vermittlung von Civic Literacy kommt der Schule und insbesondere der schulischen politischen Bildung zu (Castillo et al., 2015; Schmid, 2004). Studien verschiedener Disziplinen haben gezeigt, dass die curriculare Einbindung sowie die Qualität schulischer politischer Bildung eng mit der Civic Literacy von Jugendlichen zusammenhängt (z. B. Bayram-Özdemir et al., 2016; Bittman und Russell, 2016;

Manganelli et al., 2015; McAvoy und Hess, 2013; Schulz et al., 2018). Die Bedeutung der schulischen politischen Bildung, vor allem in der Sekundarstufe I, wird durch die Annahme weiter verstärkt, dass das Jugendalter als Lebensabschnitt, in Verbindung mit einem gleichzeitig schwindenden Einfluss des Elternhauses, als formative Phase für den Aufbau politischer Orientierungen gilt (Vollebergh et al., 2001). So deuteten Baumert et al. (2016) die Ergebnisse der BIJU Studie (Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter) als Hinweis für eine Bestätigung der „Impressionable Years“-Hypothese (siehe Visser und Krosnick, 1998), welche annimmt, dass sich politische Orientierungen vor allem im Jugendalter manifestieren und sich anschließend verfestigen.

Die schulische (politische) Bildung hat deshalb den politische formulierten Auftrag, „junge Menschen zu befähigen, sich in der modernen Gesellschaft zu orientieren und politische, gesellschaftliche und wirtschaftliche Fragen und Probleme kompetent zu beurteilen. Dabei sollen sie ermuntert werden, für Freiheit, Demokratie, Menschenrechte, Gerechtigkeit, wirtschaftliche Sicherheit und Frieden einzutreten“ (Kultusministerkonferenz, 2018, Absatz 1). Diesen Zielen seien prinzipiell alle Unterrichtsfächer verpflichtet, insbesondere die des gesellschaftswissenschaftlichen Bereichs (Ebd.). Die Rolle des Politikunterrichts, der das originäre Bezugsmedium schulischer politischer Bildung darstellt, wird hierbei allerdings nicht weiter definiert.

Eine zusätzliche Herausforderung für die schulische politische Bildung und die Beforschung dieser stellt die föderale Struktur des deutschen Bildungssystems dar. In den 16 Bundesländern gibt es mehr als 20 Bezeichnungen für Fächer der politischen Bildung – z. B.: European Studies (Bremen), Sozialkunde (u. a. Thüringen) und Gesellschaftslehre (u. a. Rheinland-Pfalz). Damit einhergeht, dass politische Bildung in den Bundesländern unterschiedlich häufig unterrichtet wird. Gökbudak und Hedtke (2020) konnten mit der systematischen Auswertung von Stundentafeln belegen, dass Lehrkräfte in der gymnasialen Sekundarstufe I in Hessen, Schleswig-Holstein und Nordrhein-Westfalen etwa acht Mal mehr Zeit haben, ihren Schülerinnen und Schülern politische Bildung nahezubringen, als Lehrkräfte in Bayern.

Zudem wird der Politikunterricht oft als das Fach beschrieben, das am häufigsten fachfremd unterrichtet wird. Tatsächlich offenbart die amtliche Schulstatistik in Nordrhein-Westfalen für das Schuljahr 2019/20, dass Politik so häufig wie kein anderes Fach von Lehrkräften ohne die entsprechende formelle Qualifikation beziehungsweise Ausbildung im Fach Politik erteilt wurde. Besonders prekär ist die Lage an Haupt- und Sekundarschulen, an denen zuletzt etwa 80 Prozent des Politikunterrichts fachfremd unterrichtet wurde (MSB, 2020, 132). Empirische Befunde deuten auf mögliche Defizite hinsichtlich der politischen Partizipation von Schülerinnen und Schülern hin. Daten der ICCS 2016 zeigten beispielsweise, dass von den in

Nordrhein-Westfalen befragten etwa 14-jährigen Schülerinnen und Schülern nur 71,7 % die Absicht haben, zukünftig an nationalen Wahlen teilzunehmen (Abs und Hahn-Laudenberg, 2017). Unter der Prämisse, dass politische Partizipation und die damit verbundene Teilnahme an politischen Wahlen eines der wichtigsten Ziele des Politikunterrichts ist, scheint dieser mit Hinblick auf die Ergebnisse der ICCS 2016 in Nordrhein-Westfalen bei einem substantiellen Anteil der Schülerinnen und Schüler nicht erfolgreich zu sein.

Die dem Kompetenzmodell zugrundeliegenden Annahmen beziehen sich in ihrer wissenschaftstheoretischen Herkunft und ihrem empirischen Anspruch auf die formale, schulische politische Bildung der Sekundarstufe I und II. Perspektivisch kann durch die theoretische und empirische Erweiterung des Modells eine Anknüpfung an andere Sozialisationsinstanzen der politischen Bildung sowie Altersgruppen gelingen. Dadurch würde es möglich werden, die Bedeutung der politischen Bildung in unterschiedlichen Kontexten empirisch einer kritischen Prüfung zu unterziehen und beispielsweise die These zu prüfen, ob sich politisches Engagement auch ohne oder sogar trotz (schulischer) politischer Bildung entwickelt.

Bedeutende Kompetenzkonzeptionen der politischen Bildung aus Theorie und Empirie

Einhergehend mit der Umorientierung von einer Input- zu einer Output-Orientierung in der deutschen Bildungslandschaft (z. B. Abs und Klieme, 2005; Klieme et al., 2003) und der Kompetenzdebatte zu Beginn der 2000er Jahre wurde zunehmend deutlich, dass ausschließlich normativ fundierte und entsprechend formulierte Anforderungen an die Schülerinnen und Schüler (allein) nicht ausreichen und darüber hinaus keine Voraussetzung für die Identifikation der Domänenstruktur innerhalb einer Civic Literacy sein können (siehe Weißeno, 2017). Stattdessen benötigt auch die outputgeleitete Kompetenzorientierung im Bereich der schulischen politischen Bildung ein empirisches Forschungsverständnis (siehe Biedermann und Reichenbach, 2009).

Einen einflussreichen, jedoch nicht unumstrittenen, Orientierungsrahmen stellt das Modell der Politikkompetenz von Detjen et al. (2012) in Weiterentwicklung des Modells der Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung von 2004 dar. Das Modell der Politikkompetenz hat einen entscheidenden Beitrag zum Wandel von einem normativ fundierten hin zu einem fachdidaktisch begründeten und kompetenzorientierten Anforderungsprofil für den Politikunterricht geleistet. Die Kritikerinnen und Kritiker des Modells, allen voran die Autorengruppe Fachdidaktik (2011), kritisieren einen zu engen inhaltlichen Zuschnitt des Modells und plädieren für eine sozialwissenschaftliche Perspektive der politischen

Bildung, mit dem Ziel der Integration verschiedener Fächer und Disziplinen (siehe Detjen, 2013).

Nach Detjen et al. (2012) kann die politische Kompetenz in die vier Dimensionen (1) Fachwissen, (2) politische Urteilsfähigkeit, (3) politische Handlungsfähigkeit sowie (4) politikbezogene Motivationen und Einstellungen gegliedert werden. Es wird hierbei ein domänenspezifischer, jedoch relativ weiter pädagogisch-psychologischer Kompetenzbegriff verfolgt, der neben politischem Wissen und Fähigkeiten auch weitere Handlungsdispositionen in den Bereichen Motivation, Einstellungen und in geringerem Ausmaße auch Volition einschließt (siehe Oberle, 2018).

Durch das Kompetenzmodell und die Diskussionen hierüber hat sich in der deutschsprachigen Politikdidaktik vor allem der Begriff der Politikkompetenz etabliert. Je nach Deutung des Begriffes der Politikkompetenz, nämlich im Sinne eines engen oder eines weiten Politikbegriffs (Sutor, 1994, S. 8f.), unterscheidet sich das der Civic Literacy zugrundeliegende Politikverständnis in seiner Intentionalität, seines Situationsbezugs und seiner Interdependenz von dem Politikverständnis der Politikkompetenz. Durch den Civic Literacy Begriff soll deutlich werden, dass ein Politikverständnis im weiteren Sinne gemeint ist. Dies bedeutet, dass auch solche Sozialgebilde, deren primärer Zweck kein politischer ist, sondern das Politische in ihnen nur situationsbedingt zum Vorschein kommt (z. B. Freundeskreis, Musikverein, Schule), betrachtet werden sollen (Ebd.).

Durch die Aufnahme weiterer Kompetenzbereiche (zusätzlich zum politischen Wissen) und der Bestrebung einer empirisch und fachdidaktisch fundierten Konzeptualisierung dieser weiteren Bereiche soll für die zukünftige Erforschung schulischer politischer Bildung ein über das Politikwissenschaftswissen allein hinausgehender Bezugsrahmen geschaffen werden. Anders als Hedtke (2011) verstehen wir Civic Literacy aber nicht als domänenüberspannende Kompetenz im Sinne einer Social Science Literacy sondern als Kompetenz, die der der Domäne Politik innewohnenden Interdisziplinarität gerecht wird und die gleichzeitig die spezifischen Anforderungen des Politikunterrichts berücksichtigt. Außerdem schließt der Begriff „Civic“ an bedeutende empirische Studien der politischen Kultur- und Sozialisationsforschung an (z. B. Civic Culture: Almond und Verba, 1963; Civics Assessment: National Assessment Governing Board, 2018; Civic and Citizenship Education Study: Schulz et al., 2016).

Für das Civics Assessment im Rahmen des US-amerikanischen National Assessment of Educational Progress (NAEP) werden die drei Dimensionen (1) Wissen, (2) Fähigkeiten und (3) Dispositionen unterschieden (National Assessment Governing Board, 2018). Die verschiedenen Dimensionen werden dabei gemeinsam über unterschiedliche Aufgabenformate (z. B.

Multiple Choice Fragen, Fragen mit kurzer oder langer Antwort) adressiert. Die jeweiligen Aufgaben bilden dabei zwei bis drei der genannten Dimensionen ab und werden anschließend zu einem Gesamtscore zusammengefasst.

Im Rahmenmodell zu der internationalen Vergleichsstudie ICCS 2016 von der International Association for the Evaluation of Educational Achievement (IEA) werden eine kognitive und eine affektiv-behaviorale Dimension definiert. Darüber hinaus wird jeder Aufgabenbereich einer von vier Inhaltsdomänen zugewiesen. Für die kognitive Dimension wurde ein Wissenstest und für die affektiv-behaviorale Dimension ein Fragebogen konzipiert und konstruiert. Die verschiedenen Teilbereiche können separat voneinander ausgewertet und interpretiert werden. Im nationalen Analyserahmen zur ICCS 2016 werden die vier Dimensionen (1) Wissen und Argumentieren, (2) Identität, (3) Partizipation sowie (4) Einstellungen und Werte unterschieden.

Zusammenfassend wird deutlich, dass alle der genannten Konzeptionen eine an der Politikwissenschaft orientierte Wissensdimension einschließen. Darüber hinaus sind jedoch auch einige Unterschiede zu beobachten. Anders als im Modell zur Politikkompetenz und der ICCS 2016 werden im Civics Assessment keine motivationalen, Einstellungs- und volitionalen Maße berücksichtigt. Behaviorale Maße in Form von bereits erfolgter Partizipation werden nur in die Konzeption des ICCS 2016 einbezogen. Anders als im fachdidaktisch begründeten Modell der Politikkompetenz vorgeschlagen, finden die Handlungs- und Urteilsfähigkeit keine Berücksichtigung in den empirischen Modellen, was vor allem in den begrenzten Möglichkeiten der Operationalisierung dieser Aspekte im Kontext der Empirischen Bildungsforschung begründet sein dürfte. Ähnlich wie bei der Handlungsfähigkeit und dem damit verbundenen politischen Handeln (siehe Detjen et al., 2012, S. 15f.) kann politisches Urteilen in der Schule nur simuliert werden. Tatsächliches Handeln und Urteilen findet aber außerhalb der schulischen politischen Bildung statt und sollte deshalb keinen Eingang in die Benotung finden. Für die Aufnahme dieser Fähigkeiten in Assessments sind besondere Anstrengungen notwendig, um ein möglichst realitätsnahes Szenario zu ermöglichen. Besonders bei größeren Stichprobengrößen kann die Erfassung von Handlungs- und Urteilsfähigkeiten deshalb schnell zu Problemen hinsichtlich der zur Verfügung stehenden Ressourcen führen. Zudem ist die Handlungsfähigkeit vor allem eine Folge von Handlungsbereitschaft und auch der Urteilsfähigkeit geht ein entsprechendes politisches Wissen voraus (Manzel und Weißeno, 2017). Folglich scheinen politische Handlungs- und Urteilsfähigkeit eher eine Folgefähigkeit der anderen Bereiche des Kompetenzmodells zu sein.

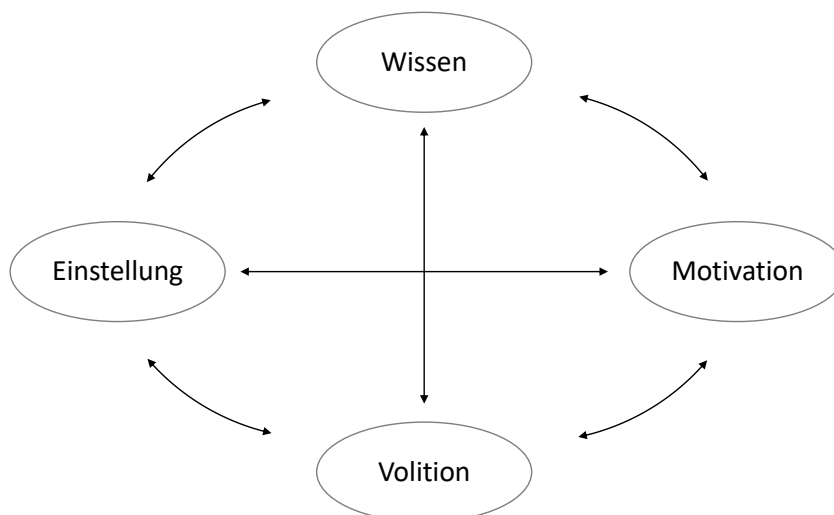
Rahmenmodell zu Civic Literacy

Den dargelegten theoretischen Überlegungen und vorgestellten Konzeptionen folgend, werden für das Rahmenmodell zu Civic Literacy ein kognitiver Bereich des politischen Wissens sowie Handlungsdispositionen in den Bereichen Motivation, Einstellung und Volition (siehe Abbildung 1) als erweiternde Bereiche abgeleitet. Neben der (empirischen) Politikdidaktik begründet sich das Rahmenmodell auch in der kognitionspsychologischen Überlegung, dass Kompetenzen „kontextspezifische Leistungsdispositionen [sind], die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme und Leutner, 2006, S. 879). Zu den Leistungsdispositionen gehören all die Dispositionen, die für die Erreichung einer bestimmten Leistung notwendig sind. Kompetenzen stellen somit eine Verknüpfung zwischen Wissen, Motivationen, Einstellungen und Volitionen her.

Im Folgenden wird die theoretische und konzeptionelle Bedeutung der verschiedenen Bereiche erläutert und diskutiert sowie für die Erfassung des politischen Wissens eine umfassende Testentwicklung vorgestellt, die eine standardisierte und über mehrere Erhebungszeitpunkte hinweg vergleichende Erfassung ermöglicht.

Abbildung 1

Vier Bereiche von Civic Literacy



Politisches Wissen

Studien, die politisches Wissen in objektiver und überprüfbarer Form erheben, sind eher selten. Ausnahmen bilden hierbei vor allem die internationale Civic and Citizenship Education Studien (Schulz et al., 2008; Schulz et al., 2016; Torney-Purta et al., 2001) und die Civics Assessment Studien des NAEP (z. B. National Assessment Governing Board, 2014, 2018) sowie die in

Deutschland durchgeführte TEESAEC-Studie (Eck und Weissenro, 2009), die POWIS-Studie (Goll et al., 2011) und die WEUS-Studie (Oberle, 2012). In allen Fällen ist der Wissenstest jedoch nur für eine Erhebung im Querschnitt konzipiert worden, sodass keine Aussagen über den (langfristigen) Wissenszuwachs gemacht werden können. Außerdem ist der Wissenstest der ICCS auf eine länderübergreifende Erhebung und somit auf länderunspezifische Inhalte und die TEESAEC- sowie die WEUS-Studie auf Wissen über die Europäische Union ausgerichtet, sodass eine Fokussierung auf Besonderheiten des politischen Systems in der Bundesrepublik Deutschland (z. B. föderales System) in diesen Studien nicht gegeben war. Aus diesem Grund wurde ein Testinstrument zur Erfassung des politischen Wissens neuentwickelt und wird im Folgenden vorgestellt.

Theorie und Konzeption des politischen Wissens

Im Kontext von Bildungsprozessen ist die Vermittlung von Wissen ein bedeutender Forschungsgegenstand. Auch für das politische Lernen und die schulische politische Bildung ist die Frage der Qualität und des Inhalts der Wissensvermittlung bedeutsam. Tatsächlich zeigen empirische Studien, dass politisches Wissen zu konsistenteren und stabileren politischen Einstellungen führt, Bürgerinnen und Bürgern beim Erreichen der eigenen Interessen hilft, die Unterstützung demokratischer Werte fördert, Vertrauen in das politische System erleichtert und zu politischer Partizipation motiviert (siehe Reichert, 2016).

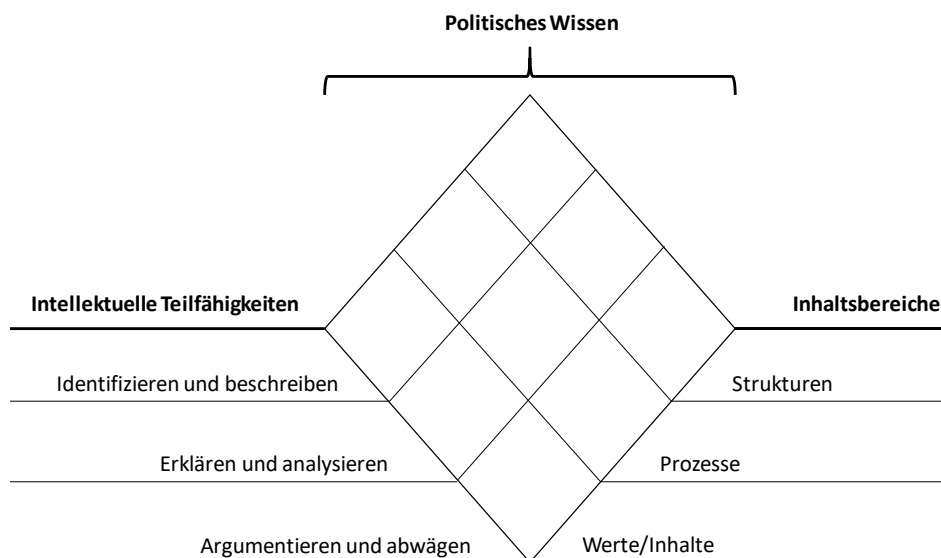
Die Wissensdimension des hier vorgestellten Rahmenmodells soll die durch die schulische politische Bildung vermittelten Wissensfacetten in den Bereichen Politik und (Zivil-)Gesellschaft adressieren. Für die Konzeption des Wissenstests wurden daher insgesamt drei Inhaltsbereiche sowie drei intellektuelle Teilfähigkeiten definiert, die in einer Matrix mit drei Zeilen und drei Spalten angeordnet werden (siehe Abbildung 2).

Die drei Inhaltsbereiche (1) Strukturen, (2) Prozesse und (3) Werte/Inhalte entsprechen weitgehend der allgemeinen Trias der Politikwissenschaft – Polity, Politics und Policy (siehe Blum und Schubert, 2018), beziehungsweise Struktur, Performanz und Werte im deutschen Kontext (siehe Fuchs, 1997) – und knüpfen an die Basiskonzepte Ordnung, Entscheidung und Gemeinwohl des politischen Fachwissens im Modell zur Politikkompetenz (siehe Weissenro et al., 2010) an. Der Inhaltsbereich „Strukturen“ befasst sich mit dem institutionellen Normengefüge und den daraus resultierenden Ordnungen und politischen Verfahren (z. B. das Parlament oder die Gewaltenteilung, siehe Schubert und Klein, 2018). Der Inhaltsbereich „Prozesse“ bezeichnet in diesem Zusammenhang die Methode der politischen Einflussnahme durch Verhandlung, Tausch, Bewältigung von Konflikten und Herstellung von Konsens (ebd.). Der

Inhaltsbereich „Werte/Inhalte“ beschreibt gesellschaftliche Grundwerte und Gegenstände des politischen Tagesgeschäfts (z. B. Freiheit, Gleichberechtigung).

Abbildung 2

Konzeption des Wissenstests zu Civic Literacy



Die drei Teilfähigkeiten (1) Identifizieren und beschreiben, (2) Erklären und analysieren und (3) Argumentieren und abwägen stellen nicht notwendigerweise aufsteigende kognitive Anforderungen dar. Die Begriffe der intellektuellen Teilfähigkeiten wurden aus dem NAEP Civics Framework abgeleitet (intellectual skills, siehe National Assessment Governing Board, 2018). Die Teilfähigkeit „Identifizieren und beschreiben“ stellt den Prozess dar, in dem Konzepten der Politik und der (Zivil-)Gesellschaft ihre Bedeutung zugewiesen und von anderen Konzepten unterschieden wird. Es soll hierdurch ein funktionales Begriffsverständnis von Politik und Zivilgesellschaft demonstriert werden, welches notwendig ist, um den vor allem sprachbasierten Grundanforderungen der Politik entsprechen und somit politischen sowie gesellschaftlichen Debatten folgen zu können. Diese Teilfähigkeit schließt damit an die zweite Anforderungsstufe des Konzepts einer modernen Allgemeinbildung von Bybee (1997) an, in der es um die bedeutungsgerechte Verwendung von Begriffen geht (siehe auch Weißeno et al., 2010, S. 19 für eine Einordnung des Konzepts in den Civic Literacy Diskurs). Die Teilfähigkeit „Erklären und analysieren“ behandelt die Verknüpfung von verschiedenen Konzepten. Die Schülerinnen und Schüler bekommen die Aufgabe, verschiedene Aspekte der Politik oder der Gesellschaft in Verbindung zueinander zu stellen und sinnvolle Argumente beziehungsweise

Erklärungen für diese Zusammenhänge zu identifizieren. Für die richtige Beantwortung dieser Items ist ein konzeptuelles Verständnis notwendig. Auch in dem Modell nach Bybee (1997) bildet das konzeptuelle Verständnis die nächste Anforderungsstufe. Schließlich beschreibt die Teilfähigkeit „Argumentieren und abwägen“ die Fähigkeit, unterschiedliche Maßnahmen für ein politisches oder gesellschaftliches Problem den jeweiligen Umständen entsprechend adäquat zu bewerten. Die Schülerinnen und Schüler müssen für die richtige Beantwortung entsprechender Fragen ein multidimensionales Politikverständnis demonstrieren. Hierfür werden die Schülerinnen und Schüler mit einer politisch oder gesellschaftlich relevanten Situation und vier unterschiedlich angemessenen Lösungsoptionen sowie einem vorgegebenen Argumentationsrahmen konfrontiert. Die Aufgabe der Schülerinnen und Schüler ist es, die entsprechend des Argumentationsrahmens beste der gegebenen Lösungsoptionen auszuwählen. Auch hier kann das Konzept einer modernen Allgemeinbildung von Bybee (1997) als Bezugsrahmen dienen, bei dem das multidimensionale Verständnis ebenfalls die höchste Anforderungsstufe darstellt. Anders als im Konzept der modernen Allgemeinbildung stellen die drei Teilfähigkeiten jedoch keine notwendigerweise hierarchisch ansteigenden Anforderungen dar.

Neuentwicklung eines Tests zum politischen Wissen

Für alle neun Zellen des Wissenstests (3 Inhaltsbereiche x 3 intellektuelle Teilfähigkeiten) wurde eine jeweils konstante Anzahl von Items konstruiert. Ein Teil der Testitems wurde aus der POWIS Studie (Goll et al., 2011) übernommen und bei Bedarf angepasst. Alle Basis- und Fachkonzepte des politischen Wissens von Weißeno et al. (2010) sind in ausgewogener Häufigkeit in den Testitems enthalten und wurden gegebenenfalls um weitere Konzepte ergänzt (z. B. Föderalismus, Partizipation und Ideologien). Die konstruierten Testitems sind dabei auf die deutsche politische Kultur zugeschnitten, ohne dabei ausschließlich auf originär für das deutsche politische System relevante Aspekte zu rekurrieren.

Das Testinstrument geht zudem über den eher engen Wissensbegriff nach Weißeno et al. (2010, S. 199) hinaus, in dem es um „das Wissen der Domäne Politik geht“. Stattdessen werden in dem hier angewandten Wissensbegriff Aspekte der Zivilgesellschaft explizit mitgedacht.

Abbildung 3

Beispiel-Items zum Wissenstest

Der Liberalismus kennzeichnet eine der großen politischen Strömungen und Ideologien.
Wie kann der Liberalismus beschrieben werden?

1 Bitte wähle eine der folgenden Antworten.

- Männer und Frauen haben die gleichen Rechte.
- Es gibt keine Grenzen und keine Regeln.
- Die freie Entfaltung des Menschen steht im Mittelpunkt.
- Jeder ist für sich selbst verantwortlich.

Warum kann es sinnvoll sein, Gesetze auf internationaler, anstatt auf nationaler, Ebene zu beschließen?
Internationale Gesetze ...

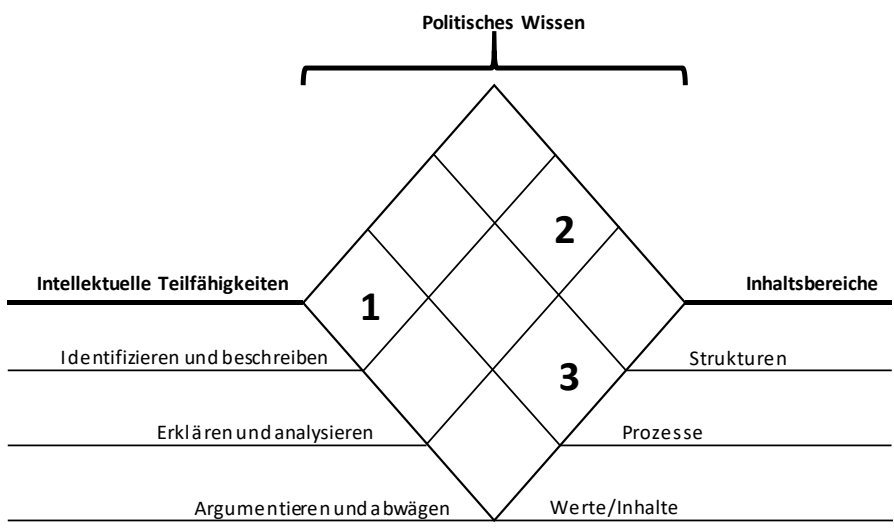
2 Bitte wähle eine der folgenden Antworten.

- können die Zusammenarbeit zwischen Ländern erleichtern, weil dann die gleichen Gesetze für alle gelten.
- ermöglichen es, unangenehme nationale Gesetze zu umgehen.
- müssen seltener neu verhandelt werden, weil sie auf höherer Ebene beschlossen wurden.
- sind wichtiger als nationale Gesetze und werden deshalb häufiger eingehalten.

Du möchtest Schulsprecherin bzw. Schulsprecher werden. Allerdings möchte auch Klara Schulsprecherin werden.
Wie solltest du, im Sinne eines demokratischen Wahlkampfes, am besten vorgehen?

3 Bitte wähle eine der folgenden Antworten.

- Du informierst die Schülerinnen und Schüler darüber, was du machen möchtest, wenn du gewählt wirst.
- Du verteilst bei Kindern aus unteren Klassenstufen Bonbons; bei Älteren verlost du Gutscheine.
- Du lässt Klara den Vortritt. Es wäre schlecht, wenn es mehr als eine Kandidatin oder einen Kandidaten gibt.
- Du bittest Klara darum, nicht gegen dich anzutreten, weil du schon so lange Schulsprecherin bzw. Schulsprecher werden möchtest.



Es wurden Testitems mit Fokus auf die Klassenstufen 7 und 10 konstruiert, da davon ausgegangen werden darf, dass sich vor allem in dieser Altersspanne politisches Wissen manifestiert und politische Orientierungen formieren und verfestigen (siehe Baumert et al., 2016). Die Testitems wurden als geschlossene Fragen im Single-Choice-Format mit drei Distraktoren, die möglichst typische Prä- und Misskonzepte repräsentieren, konstruiert (siehe Abbildung 3) und entsprechen methodisch den gängigen Itemkonstruktionsrichtlinien (Haladyna et al., 2002).

Alle potentiellen Items des Wissenstests wurden durch ein Gremium bestehend aus drei Professoren und einer Professorin der deutschen Politikdidaktik sowie einer Projektleitung einer Studie zu Civic Literacy auf Einhaltung der inhaltlichen und methodischen Itemkonstruktionsrichtlinien überprüft und hinsichtlich ihrer Verortung im Wissenstest geratet. Die Beratung sowie das Rating durch das Expertengremium erfolgten in zwei Phasen. In der ersten Phase wurde das vorgeschlagene Modell mit den Mitgliedern des Gremiums diskutiert und eine Teilmenge der konstruierten Testitems hinsichtlich Relevanz und vorläufiger Verortung in dem Modell bewertet. Auf Basis dieser Rückmeldungen wurden neue Testitems konstruiert und bereits existierende überarbeitet. In der zweiten Phase wurden alle 143 konstruierten Testitems durch die Expertin und Experten hinsichtlich Relevanz, Schwierigkeit und Verortung im Modell des Wissenstests geratet, sodass jedes Testitem mindestens zwei Mal, in den meisten Fällen häufiger, geratet wurde. Auf Basis der Ergebnisse zu Relevanz und Schwierigkeit wurden 99 Testitems ausgewählt und den Klassenstufen 7 und 10 zugeordnet. Neben jeweils 36 klassenstufenspezifischen für die Klassenstufen 7 beziehungsweise 10 sind auch 27 klassenstufenunspezifische Items identifiziert worden, sodass insgesamt 63 Items pro Klassenstufe eingesetzt werden können. Durch die klassenstufenunspezifischen Items wird die Möglichkeit eines Linking gegeben, damit auch Wissenszuwächse quantifiziert und weitergehend analysiert werden können. Alle neun Teilbereiche des Modells zum Wissenstest werden in gleichbleibender Testitemanzahl abgebildet. Inwieweit die Expertenratings mit den zuvor durch die Itemkonstrukteure und -konstrukteurinnen vorgenommenen Verortungen der einzelnen Items im Modell des Wissenstests übereinstimmen, wird für die 99 ausgewählten Testitems in der Abbildung 4 dargestellt.

Abbildung 4*Ergebnisse des Expertenratings*

		Inhaltsbereiche			Übereinstimmung Fähigkeit	
		Strukturen	Prozesse	Werte/Inhalte	K7	K10
Intellektuelle Fähigkeiten						
Identifizieren und beschreiben		84%	82%	65 Ratings	55 Ratings	
Erklären und analysieren		75%	72%	52 Ratings	64 Ratings	
Urteilen, Positionen beziehen und verteidigen		66%	66%	53 Ratings	56 Ratings	
		83%	76%	90%		
Übereinstimmung Inhaltsbereich	K7	54 Ratings	59 Ratings	57 Ratings	Gesamt: 78% 340 Ratings	
	K10	75%	72%	76%	Gesamt: 74% 350 Ratings	
		55 Ratings	54 Ratings	66 Ratings		

Insgesamt wurde der jeweilige Inhaltsbereich beziehungsweise die jeweilige intellektuelle Teilfähigkeit für 78 % aller Testitems der siebten und für 74 % aller Testitems der zehnten Klassenstufe richtig geratet. Der Anteil richtig gerateter Testitems weicht somit deutlich positiv von einer randomisierten Verteilung der Ratings (33 %) ab. Für jede intellektuelle Fähigkeit und jeden Inhaltsbereich wurden jeweils 21 Items für die Klassenstufen 7 und 10 geratet. Zudem hatten die Mitglieder des Gremiums in der zweiten Phase erneut die Möglichkeit, Testitems mit Kommentaren zu versehen, welche zur weiteren Optimierung der Testitems genutzt wurden. Anschließend wurden die Testitems von Schülerinnen und Schülern verschiedener Klassenstufen und Schulformen in kognitiven Prätests bearbeitet, auf ihre sprachliche sowie inhaltliche Schwierigkeit geprüft und bei Bedarf erneut angepasst. Die hierbei notwendig gewordenen Anpassungen waren jedoch von sehr geringem Ausmaß. Perspektivisch kann die Gesamtanzahl der Testitems bei Bedarf und unter Beachtung der Verteilung im Modell weiter reduziert werden. Die Verortung des Wissenstests im Gesamtmodell wird in Abbildung 5 dargestellt.

Theorie und Konzeption politischer Motivationen

In der Kognitionspsychologie wird Motivation als das antizipierte Produkt einer Interaktion von Person und Situation verstanden (Heckhausen und Heckhausen, 2010). Das bedeutet, dass die

in einer bestimmten Situation „vorhandene Motivation einer Person, ein bestimmtes Ziel anzustreben, [...] von personenbezogenen und von situationsbezogenen Einflüssen geprägt“ wird (ebd., S. 3). Für die personenbezogenen Einflüsse kann zwischen universellen Verhaltenstendenzen in Form von elementaren psychischen Bedürfnissen und impliziten sowie expliziten Motiven unterschieden werden. Für die Kompetenzforschung sind vor allem die expliziten Motive von Interesse, da diese (1) im Gegensatz zu den universellen Bedürfnissen domänenspezifisch sein können und (2) anders als implizite Motive den Methoden des Selbstberichts zugänglich sind (Brunstein, 2010). „Explizite Motive spiegeln [...] die Selbstbilder, Werte und Ziele wider, die sich eine Person selbst zuschreibt und mit denen sie sich identifiziert“ (ebd., S. 237). In der politikdidaktischen Forschung zu politischen Motivationen sind vor allem das (habituelle) politische Interesse sowie das politische Effektivitätsgefühl bedeutsam (Chrysochoou und Barrett, 2017; Oberle, 2018; Rasmussen und Nørgaard, 2018).

Das politische Interesse wird für gewöhnlich als Vorbedingung demokratischer Teilhabe (democratic citizenship) verstanden (Neundorf et al., 2013). Verschiedene Studien zeigten einen direkten Zusammenhang zwischen politischem Interesse und politischer Partizipation (z. B. Pavlopoulos et al., 2019; Wang et al., 2019). Das habituelle beziehungsweise überdauernde und persönliche Interesse ist „eine anhaltende individuelle Präferenz“, die sich zunehmend stabilisiert und mit einem positiven Affekt in Bezug auf das Interessensobjekt einhergeht (Landwehr, 2017a, S. 94; siehe auch Prior und Bougher, 2018). Es ist somit weniger situationsbeziehungsweise aufgabenabhängig als das situationsspezifische Interesse. Für die Beschreibung des politischen Interesses unterscheiden Detjen et al. (2012) folgende drei Aspekte: (1) selektive Aufmerksamkeit gegenüber politischen Objekten und Ereignissen, (2) relativer Bedeutungsüberschuss von Politik gegenüber anderen Lebensbereichen und (3) das Bestreben, den Gegenstandsbereich Politik verstehen zu können (siehe Russo und Stattin, 2017).

Die Selbstwirksamkeit, beziehungsweise das individuell empfundene Effektivitätsgefühl, kann als Überzeugung eines Menschen verstanden werden, auch herausfordernde Situationen selbständig oder kollektiv bewältigen zu können (Bandura, 1977). Das politische Effektivitätsgefühl kann als die Einschätzung der eigenen oder kollektiven Effektivität im Bereich der Politik beschrieben werden (Oberle, 2018). Für das politische Effektivitätsgefühl kann zwischen dem internen und dem externen politischen Effektivitätsgefühl unterschieden werden (z. B. Rasmussen und Nørgaard, 2018). Während das interne politische Effektivitätsgefühl die eigene Fähigkeit zur Beeinflussung der politischen Willensbildung (gelegentlich auch beschrieben als politische Selbstwirksamkeit, siehe Oberle und Wenzel, 2019) ausdrückt, beschreibt das

externe politische Effektivitätsgefühl die wahrgenommene allgemeine Responsivität beziehungsweise kollektive Beeinflussbarkeit des politischen Systems.

Das politische Interesse sowie das politische Effektivitätsgefühl sind demnach zentrale Facetten der politischen Motivation und somit Teil des Kompetenzmodells zu Civic Literacy.

Theorie und Konzeption politischer Einstellungen

Einstellungen sind psychische Tendenzen, die dadurch zum Ausdruck kommen, dass ein bestimmtes Objekt mit einem gewissen Grad an Zuneigung oder Ablehnung bewertet wird (Eagly und Chaiken, 1993), wofür auf kognitive, affektive oder behaviorale Informationen zurückgegriffen werden kann (Maio et al., 2018). Einstellungen werden folglich durch wertende Urteile zum Ausdruck gebracht. Dabei unterscheiden sich Einstellungen einerseits in ihrer Valenz – zum Beispiel kann ein Objekt positiv, neutral oder negativ bewertet werden – und andererseits in ihrer Intensität (ebd.).

Detjen et al. (2012, S. 92) nehmen explizit Bezug auf politische Einstellungen und bezeichnen diese als „erworbene, dauerhafte, häufig stärker gefühlsbezogene Orientierungen gegenüber Klassen von Personen oder Objekten, die wiederum wichtige Prädispositionen für tatsächliches Handeln darstellen“. Zu unterscheiden sind politische Einstellungen von anderen Einstellungen durch ihren Bezug auf politische Sachverhalte (siehe Gabriel, 1994).

Die Einstellung gegenüber (staatsrelevanten) Objekten kann durch die Einstellung zur Demokratie abgebildet werden. Für die Einstellung zur Demokratie ist vor allem zwischen der prinzipiellen Zustimmung zu der Idee der Demokratie und ihren konstituierenden Prinzipien einerseits und der Bewertung des aktuellen Zustandes der Demokratie andererseits zu unterscheiden (Gabriel, 2000). Um eine Vergleichbarkeit über verschiedene Messzeitpunkte und Regionen hinweg zu ermöglichen, wird im Kompetenzmodell auf die Einstellung zur prinzipiellen Idee der Demokratie fokussiert.

Einstellungen hinsichtlich gesellschaftlicher Gruppen beziehungsweise Klassen von Personen können über die Einstellung zur pluralen Gesellschaft betrachtet werden. Das Plurale ist neben dem Liberalen von konstituierender Bedeutung für die deutsche Demokratie und gewinnt als Gegenstand der politischen Bildung vor allem vor dem Hintergrund einer zunehmend polarisierten Gesellschaft an Bedeutung (Albert et al., 2019; Rau und Stier, 2019).

Relevante Einstellungsfacetten im Kompetenzmodell zu Civic Literacy sind daher die Einstellung zur Demokratie sowie die Einstellung zur pluralen Gesellschaft.

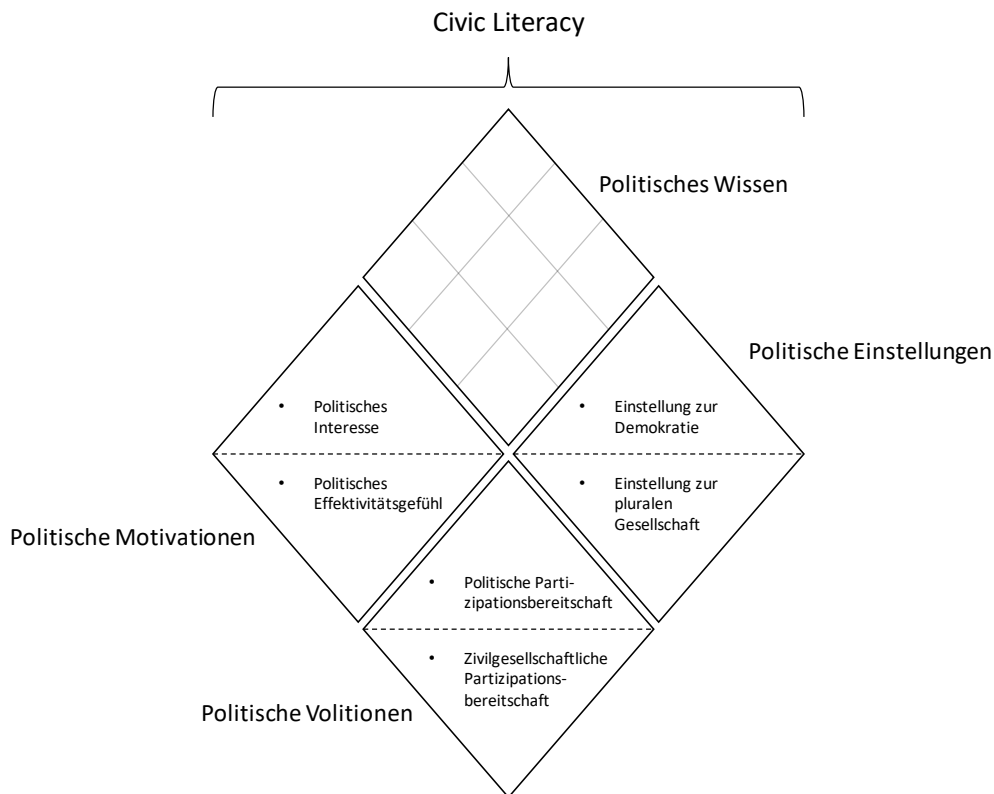
Theorie und Konzeption politischer Volitionen

Volitionen werden im Allgemeinen als regulative Prozesse verstanden, die darüber entscheiden, welche Motivationstendenzen bei welchen Gelegenheiten und auf welche Weise realisiert werden sollen (Heckhausen und Heckhausen, 2010). Es handelt sich dabei um kognitive Prozesse, die im Rahmen der präaktionalen und aktionalen Handlungsphase auf die konkrete Realisierung von Zielen im Handeln bezogen sind (Achtziger und Gollwitzer, 2010). Für die Formation von Handlungsabsichten wirken „attitude toward the behavior, subjective norm, and perception of behavioral control“ zusammen (Ajzen, 1991, 2002, S. 665).

Im Kontext der politischen Bildung wird unter Volition die Partizipationsbereitschaft auf politischer und zivilgesellschaftlicher Ebene zusammengefasst. Auch für politisches und zivilgesellschaftliches Handeln gilt, dass „die Handlungsinitiierung erfolgt, wenn die Volitionsstärke der Zielintention ausreichend ist, das heißt die Person sich stark genug der Zielerreichung verpflichtet fühlt“ (Landwehr, 2017b, S. 87). Die Intensität und Ausdauer der initiierten Handlung hängt dabei von der Volitionsstärke der Zielintention ab (ebd.). Demnach ist festzuhalten, dass die Partizipationsbereitschaft selbst im Zielbereich (auch schulischer) politischer Bildung liegt und als wünschenswertes Resultat definiert werden kann (Oberle, 2018).

Die volitionalen Dispositionen können beispielsweise adressiert werden, indem Schülerinnen und Schüler zu ihren Intentionen hinsichtlich der Initiierung wichtiger politischer und zivilgesellschaftlicher Zielhandlungen befragt werden. Für die Unterscheidung zwischen politischen und zivilgesellschaftlichen Zielhandlungen kann auf das Modell zur politischen Teilhabe von Ekman und Amnå (2012) zurückgegriffen werden. Während für die politischen Zielhandlungen (manifest political participation) zwischen formaler politischer Partizipation und politischem Aktivismus unterschieden werden kann, gliedert sich die zivilgesellschaftliche Partizipation (civil participation, bzw. latent political participation) in politische Aufmerksamkeit und ziviles Engagement.

Beide Facetten der politischen Volition – die politische Partizipationsbereitschaft wie auch die zivilgesellschaftliche Partizipationsbereitschaft – sind integraler Bestandteil des vorgestellten Rahmenmodells zu Civic Literacy.

Abbildung 5*Rahmenmodell zu Civic Literacy***Resümee und Ausblick**

Spätestens die überaus populäre und weltweit expandierende soziale Bewegung „Fridays-for-Future“ hat offenbart, dass eine aktive Teilnahme junger Menschen an der (Zivil-)Gesellschaft, erheblichen gesellschaftlichen Wandel anstoßen kann. Nichtsdestotrotz zeigen die eingangs erwähnten Vorfälle im US-amerikanischen Kapitol im Januar 2021 sowie vor dem deutschen Bundestag wenige Monate zuvor, dass auch demokratiefeindliche politische Bewegungen von einer leichteren und anonymen Vernetzung durch die neuen sozialen Medien profitieren.

Umso wichtiger erscheint es, dass auch junge Menschen bereits ein Verständnis unseres politischen und (zivil-)gesellschaftlichen Systems haben und der Genese etwaiger demokratiefeindlicher oder -bedrohlicher Einstellungen vorgebeugt wird. Eine bedeutende Rolle kann dabei der schulischen politischen Bildung zukommen.

Obwohl das Jugend- und das junge Erwachsenenalter als allgemein kritische Phase in der Entwicklung, Aufrechterhaltung und Manifestation von Civic Literacy erachtet wird, liegen empirische Belege bislang nur für Teilabschnitte vor. Längsschnittliche Studien mit kurzen Messintervallen über einen längeren Zeitraum, die neben der Kompetenzentwicklung auch die

Betrachtung verschiedener Bedingungsfaktoren ermöglichen, fehlen noch weitgehend (für eine Übersicht siehe Grob, 2009; und Kuger und Gille, 2020). Da Civic Literacy bei Jugendlichen, vor allem im Längsschnitt, noch wenig beforscht wurde, fehlten bisher auch standardisierte und vergleichbare Messinstrumente, die für den Einsatz in verschiedenen Altersstufen geeignet und verlinkt sind. Dies betrifft vor allem den hier vorgestellten Wissenstest.

Ein besonderes Desiderat für die zukünftige Kompetenzforschung zu Civic Literacy ist es, mehr über die Rolle der verschiedenen Sozialisationsinstanzen zu erfahren. Hierbei sind vor allem die Schule und der Politikunterricht, der familiäre Hintergrund sowie die Peers der interessierenden Schülerinnen und Schüler von Bedeutung. Eine besondere Herausforderung wird dabei sein, die verschiedenen schul- und bundeslandspezifischen Vorbedingungen (z. B. Umfang des Politikunterrichts, Integration in andere Fächer, fachfremder Unterricht) in angemessener Weise zu beachten und zu erfassen. Die genaue Betrachtung dieser Vorbedingungen und der verschiedenen Bedingungsfaktoren von Civic Literacy ist bedeutsam für den weiteren Erkenntnisgewinn.

Das hier skizzierte Rahmenmodell zur empirischen Erfassung von Civic Literacy mit seinen vier Kompetenzbereichen soll nach einer empirischen Erprobung zukünftig in Studien der politischen Kultur- und Sozialisationsforschung sowie der Empirischen Bildungsforschung und der politischen Psychologie Anwendung finden. Hierzu gehört insbesondere das Nationale Bildungspanel (NEPS; Blossfeld und Roßbach, 2019), für das Civic Literacy in den Kanon der zu erfassenden Kompetenzen in den Klassenstufen 7 und 10 aufgenommen wurde. Dadurch können zukünftig wichtige Erkenntnisse für die Erforschung und Praxis der (schulischen) politischen Bildung im Jugendalter gewonnen werden.

Literatur

- Abs, H. J. & Hahn-Laudenberg, K. (2017). *Das politische Mindset von 14-Jährigen: Ergebnisse der International Civic and Citizenship Education Study 2016*. Waxmann.
- Abs, H. J. & Klieme, E. (2005). Standards für schulbezogene Evaluation. *Zeitschrift für Erziehungswissenschaft*, 4, 45–62. https://doi.org/10.1007/978-3-322-80769-4_4
- Achour, S. & Wagner, S. (2019). *Wer hat, dem wird gegeben: Politische Bildung an Schulen: Bestandsaufnahme, Rückschlüsse und Handlungsempfehlungen* (Schriftenreihe des Netzwerk Bildung). Friedrich-Ebert-Stiftung.
- Achtziger, A. & Gollwitzer, P. M. (2010). Motivation und Volition im Handlungsverlauf. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (4. Aufl., S. 309–335). Springer. https://doi.org/10.1007/978-3-642-12693-2_12
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology*, 32(4), 665–683. <https://doi.org/10.1111/j.1559-1816.2002.tb00236.x>
- Albert, M., Hurrelmann, K., Quenzel, G. & Kantar (Hrsg.). (2019). *18. Shell Jugendstudie: Jugend 2019 - eine Generation meldet sich zu Wort*. Beltz.
- Almond, G. A. & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.
- Autorengruppe Fachdidaktik (Hrsg.). (2011). *Politik und Bildung: Bd. 64. Konzepte der politischen Bildung: Eine Streitschrift*. Wochenschau Verlag.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)
- Baumert, J., Becker, M., Cortina, K. S., Köller, O., Kropf, M. & Maaz, K. (2016). Die Entwicklung des politischen Interesses und des Selbstkonzepts der politischen Kompetenz vom Jugend- bis in das Erwachsenenalter. In C. Grunert, A. Schippling & N. Pfaff (Hrsg.), *Kritische Bildungsforschung: Standortbestimmungen und Gegenstandsfelder* (S. 323–352). Barbara Budrich.
- Bayram-Özdemir, S., Stattin, H. & Özdemir, M. (2016). Youth's initiations of civic and political discussions in class: Do youth's perceptions of teachers' behaviors matter and

- why? *Journal of Youth and Adolescence*, 45(11), 2233–2245.
<https://doi.org/10.1007/s10964-016-0525-z>
- Biedermann, H. & Reichenbach, R. (2009). Die empirische Erforschung der politischen Bildung und das Konzept der politischen Urteilskompetenz. *Zeitschrift für Pädagogik*, 55(6), 872–886.
- Bittman, B. & Russell, W. (2016). Civic education in United States: A multiple regression of civic education scores from the National Assessment of Educational Progress. *Research in Social Sciences and Technology*, 1(2), 1–16.
<https://doi.org/10.46303/ressat.01.02.1>
- Blossfeld, H.-P. & Roßbach, H.-G. (2019). *Education as a Lifelong Process: The German National Educational Panel Study (NEPS)* (2. Aufl.). Springer.
<https://doi.org/10.1007/978-3-658-23162-0>
- Blum, S. & Schubert, K. (2018). *Politikfeldanalyse: Eine Einführung*: Springer.
<https://doi.org/10.1007/978-3-658-17758-4>
- Brunstein, J. C. (2010). Implizite und explizite Motive. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (4. Aufl., S. 237–255). Springer:
https://doi.org/10.1007/978-3-642-12693-2_9
- Bybee, R. W. (1997). *Towards an understanding of scientific literacy*. In *Scientific Literacy: An international symposium, 1997*. IPN Kiel.
- Castillo, J. C., Miranda, D., Bonhomme, M., Cox, C. & Bascopé, M. (2015). Mitigating the political participation gap from the school: the roles of civic knowledge and classroom climate. *Journal of Youth Studies*, 18(1), 16–35.
<https://doi.org/10.1080/13676261.2014.933199>
- Chrysoschoou, X. & Barrett, M. (2017). Civic and political engagement in youth: Findings and prospects. *Zeitschrift für Psychologie*, 225(4), 291–301.
- Detjen, J. (2013). *Politische Bildung: Geschichte und Gegenwart in Deutschland* (2. Aufl.). Lehr- und Handbücher der Politikwissenschaft. Oldenbourg Wissenschaftsverlag Verlag. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=757539> <https://doi.org/10.1524/9783486741902>
- Detjen, J., Massing, P., Richter, D. & Weißeno, G. (2012). *Politikkompetenz – ein Modell*. Springer.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes* (2. Aufl.). Harcourt Brace Jovanovich.

- Eck, V. & Weisseno, G. (2009). Political knowledge of 14- to 15-year-old students: Results of the TEESAEC intervention study in Germany. In G. Weißeno & V. Eck (Hrsg.), *Teaching European Citizens: A quasi-experimental study in six countries* (S. 19–30). Waxmann.
- Edelstein, W. & Fauser, P. (2001). *"Demokratie lernen und leben" Gutachten für ein Modellversuchsprogramm der BLK*. Bonn. <http://www.blk-bonn.de/papers/heft96.pdf>
- Ekman, J. & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3), 283–300. <https://doi.org/10.2478/s13374-012-0024-1>
- Fuchs, D. (1997). Welche Demokratie wollen die Deutschen? In O. W. Gabriel (Hrsg.), *Politische Orientierungen und Verhaltensweisen im vereinigten Deutschland* (S. 81–113). Springer. https://doi.org/10.1007/978-3-322-95869-3_4
- Gabriel, O. W. (1994). Politische Einstellungen und politische Kultur - zur Klärung der Begriffe und ihrer politikwissenschaftlichen Bedeutung. In O. W. Gabriel (Hrsg.), *Die EU-Staaten im Vergleich: Strukturen, Prozesse, Politikinhalt* (2. Aufl., S. 96–131). Westdt. Verl. <https://doi.org/10.18419/OPUS-5523>
- Gabriel, O. W. (2000). Demokratische Einstellungen in einem Land ohne demokratische Traditionen? Die Unterstützung der Demokratie in den neuen Bundesländern im Ost-West-Vergleich. In J. W. Falter, O. W. Gabriel & H. Rattinger (Hrsg.), *Wirklich ein Volk? Die politischen Orientierungen von Ost- und Westdeutschen im Vergleich* (S. 41–77). Springer. https://doi.org/10.1007/978-3-663-05704-8_2
- Gökbudak, M. & Hedtke, R. [R.]. (2020). *3. Ranking Politische Bildung: Politische Bildung an allgemeinbildenden Schulen der Sekundarstufe I im Bundesländervergleich 2019* (Didaktik der Sozialwissenschaften: Working Papers Nr. 11). Bielefeld. Fakultät für Soziologie. https://pub.uni-bielefeld.de/download/2941780/2943402/Ranking_Politische_Bildung_2019_final_1.pdf
- Goll, T., Richter, D. & Weißeno, G. (2011). Politisches Wissen von Schüler/-innen mit und ohne Migrationshintergrund (POWIS): Ergebnisse einer Studie. In D. Lange (Hrsg.), *Entgrenzungen: Gesellschaftlicher Wandel und politische Bildung* (S. 126–131). Wochenschau Verlag.
- Grob, U. (2009). Die Entwicklung politischer Orientierung vom Jugend- ins Erwachsenenalter: - Ist die Jugend eine spezifisch sensible Phase in der politischen Sozialisation? In H. Fend, U. Grob & F. Berger (Hrsg.), *Lebensverläufe, Lebensbewältigung, Lebensglück: Ergebnisse der Life-Studie* (S. 305–326). Springer.

- Hahn-Laudenberg, K. (2020). Politisches Wissen von Schüler innen über Grund rechte und das parlamentarische Regierungssystem – Herausforderungen für die schulische Auseinandersetzung mit zentralen Inhalten des Grundgesetzes. In T. Goll & B. Minkau (Hrsg.), *Verfassung und Verfassungsrecht als Gegenstand politischer Bildung. Das Grundgesetz für die Bundesrepublik Deutschland* (S. 51–70). Barbara Budrich. <https://doi.org/10.2307/j.ctv12sdvs6.8>
- Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–333. https://doi.org/10.1207/S15324818AME1503_5
- Heckhausen, J. & Heckhausen, H. (2010). *Motivation und Handeln* (4. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-12693-2>
- Hedtke, R. (2011). Die politische Domäne im sozialwissenschaftlichen Feld. In Autorengruppe Fachdidaktik (Hrsg.), *Politik und Bildung: Bd. 64. Konzepte der politischen Bildung: Eine Streitschrift*. Wochenschau Verlag.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., & Tenorth, H.-E. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. BMBF.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen: Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876–903.
- Kuger, S. & Gille, M. (2020). Entwicklung des zivilgesellschaftlichen Engagements im Jugend- und jungen Erwachsenenalter. *Zeitschrift für Erziehungswissenschaft*, 23(6), 1103–1123. <https://doi.org/10.1007/s11618-020-00978-2>
- Kultusministerkonferenz. (2018). *Demokratiebildung*. <https://www.kmk.org/themen/allgemeinbildende-schulen/weitere-unterrichtsinhalte/demokratiebildung.html>
- Landwehr, B. (2017a). *Partizipation, Wissen und Motivation im Politikunterricht*. Springer. <https://doi.org/10.1007/978-3-658-16507-9>
- Landwehr, B. (2017b). Politische Motivation. In B. Landwehr (Hrsg.), *Partizipation, Wissen und Motivation im Politikunterricht* (S. 69–108). Springer. https://doi.org/10.1007/978-3-658-16507-9_4
- Maio, G. R., Haddock, G. & Verplanken, B. (2018). *The psychology of attitudes and attitude change* (3. Aufl.). Sage Publications.

- Manganelli, S., Lucidi, F. & Alivernini, F. (2015). Italian adolescents' civic engagement and open classroom climate: The mediating role of self-efficacy. *Journal of Applied Developmental Psychology, 41*, 8–18. <https://doi.org/10.1016/j.appdev.2015.07.001>
- Manzel, S. & Weißeno, G. (2017). Modell der politischen Urteilsfähigkeit – eine Dimension der Politikkompetenz. In M. Oberle & G. Weißeno (Hrsg.), *Politikwissenschaft und Politikdidaktik: Theorie und Empirie* (S. 59–86). Springer. https://doi.org/10.1007/978-3-658-07246-9_5
- McAvoy, P. & Hess, D. (2013). Classroom Deliberation in an Era of Political Polarization. *Curriculum Inquiry, 43*(1), 14–47. <https://doi.org/10.1111/curi.12000>
- Milner, H. (2002). *Civic Literacy: How Informed Citizens Make Democracy Work*. University Press of New England. <https://books.google.de/books?id=zy-2GAAAAMAAJ>
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2020). *Das Schulwesen in Nordrhein-Westfalen aus quantitativer Sicht 2019/20.: Statistische Übersicht Nr. 408*. Düsseldorf. https://www.schulministerium.nrw/system/-files/media/document/file/quantita_2019.pdf
- National Assessment Governing Board. (2014). *Civics framework for the 2014 National Assessment of Educational Progress*. U.S. Department of Education. <https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/civics/2014-civics-framework.pdf>
- National Assessment Governing Board. (2018). *Civics Framework: for the 2018 National Assessment of Educational Progress*. U.S. Department of Education. <https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/civics/2018-civics-framework.pdf>
- Neundorf, A., Smets, K. & García-Albacete, G. M. (2013). Homemade citizens: The development of political interest during adolescence and young adulthood. *Acta Politica, 48*(1), 92–116. <https://doi.org/10.1057/ap.2012.23>
- Oberle, M. (2012). *Politisches Wissen über die Europäische Union: Subjektive und objektive Politikkenntnisse von Jugendlichen*. Springer. <https://doi.org/10.1007/978-3-531-94171-4>
- Oberle, M. (2018). Politisches Effektivitätsgefühl von Schüler/-innen. In S. Manzel & M. Oberle (Hrsg.), *Kompetenzorientierung* (S. 85–97). Springer. https://doi.org/10.1007/978-3-658-16889-6_8

- Oberle, M. & Wenzel, N. (2019). Politisches Vertrauen und Effektivitätsgefühl von Schüler/innen - Einflussfaktoren und Relevanz für politische Partizipation. In M. Lotz & K. Pohl (Hrsg.), *Schriftenreihe der GPJE. Gesellschaft im Wandel: Neue Aufgaben für die politische Bildung und ihre Didaktik* (S. 176–186). Wochenschau Verlag.
- Partnership for 21st Century Skills. (2009). *P21 framework definitions*. <https://files.eric.ed.gov/fulltext/ED519462.pdf>
- Pavlopoulos, V., Kostoglou, D. & Motti-Stefanidi, F. (2019). From political interest to participation in EU-related actions: The mediating role of European identity and political efficacy. *Psychology: The Journal of the Hellenic Psychological Society*, 24(2), 102–121.
- Prior, M. & Bougher, L. D. (2018). “Like they’ve never, ever seen in this country”? Political interest and voter engagement in 2016. *Public Opinion Quarterly*, 82(S1), 822–842.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
- Rasmussen, S. H. R. & Nørgaard, A. S. (2018). When and why does education matter? Motivation and resource effects in political efficacy. *European Journal of Political Research*, 57(1), 24–46.
- Rau, J. P. & Stier, S. (2019). Die Echokammer-Hypothese: Fragmentierung der Öffentlichkeit und politische Polarisierung durch digitale Medien? *Zeitschrift für Vergleichende Politikwissenschaft*, 13(3), 399–417. <https://doi.org/10.1007/s12286-019-00429-1>
- Reichert, F. (2016). How Internal Political Efficacy Translates Political Knowledge Into Political Participation: Evidence From Germany. *Europe's Journal of Psychology*, 12(2), 221–241. <https://doi.org/10.5964/ejop.v12i2.1095>
- Russo, S. & Stattin, H. (2017). Stability and change in youths’ political interest. *Social indicators research*, 132(2), 643–658.
- Schmid, C. (2004). *Politisches Interesse von Jugendlichen: Eine Längsschnittuntersuchung zum Einfluss von Eltern, Gleichaltrigen, Massenmedien und Schulunterricht*. Deutscher Universitäts-Verlag. <http://dx.doi.org/10.1007/978-3-322-81338-1>
- Schubert, K. & Klein, M. (2018). *Das Politiklexikon: Begriffe, Fakten, Zusammenhänge* (7. Aufl.). Dietz.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B. & Agrusti, G. (2016). *IEA International Civic and Citizenship Education Study 2016: Assessment Framework*. Springer.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G. & Friedman, T. (2018). *Becoming Citizens in a Changing World: IEA International Civic and Citizenship Education*

- Study 2016: International Report*. Springer. <https://doi.org/10.1007/978-3-319-73963-2>
- Schulz, W., Fraillon, J., Ainley, J., Losito, B. & Kerr, D. (2008). *International Civic and Citizenship Education Study: Assessment Framework*. Springer.
- Sutor, B. (1994). *Politische Bildung als Praxis: Grundzüge eines didaktischen Konzepts* (2. Aufl.). Politische Bildung Kleine Reihe: Bd. 7. Wochenschau Verlag.
- Tocqueville, A. de. (1838). *Democracy in America*. G. Dearborn & Co.
- Torney-Purta, J., Lehmann, R., Oswald, H. & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at age fourteen*. The International Association for the Evaluation of Educational Achievement. <http://www.wam.umd.edu/%7Eiea/>
- Visser, P. S. & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology*, 75(6), 1389–1410. <https://doi.org/10.1037/0022-3514.75.6.1389>
- Vollebergh, W. A. M., Iedema, J. & Raaijmakers, Q. A. (2001). Intergenerational Transmission and the Formation of Cultural Orientations in Adolescence and Young Adulthood. *Journal of Marriage and Family*, 63(4), 1185–1198. <http://www.jstor.org/stable/3599823>
- Wang, C.-H., Weng, D. L.-C. & Tsai, C. (2019). Personality traits and political participation in Taiwan: a mediation approach. *Political Science*, 71(3), 175–192. <https://doi.org/10.1080/00323187.2020.1767506>
- Weißeno, G. (2017). Politikdidaktische Theoriebildung – eine wissenschaftstheoretische Orientierung. In M. Oberle & G. Weißeno (Hrsg.), *Politische Bildung. Politikwissenschaft und Politikdidaktik: Theorie und Empirie* (S. 1–16). Springer. https://doi.org/10.1007/978-3-658-07246-9_1
- Weißeno, G., Detjen, J., Juchler, I., Massing, P. & Richter, D. (2010). *Konzepte der Politik - ein Kompetenzmodell*. Schriftenreihe: Bd. 1016. Bundeszentrale für politische Bildung. https://www.pedocs.de/volltexte/2016/12009/pdf/Weisseno-et_al_2010_Konzepte-_der_Politik_.pdf

2.2 Towards the Measurability of a Competence Model for School-based Civic Education

Alscher, P., Grecu, A. L., Ludewig, U. & McElvany, N. (2023). Towards the measurability of a competence model for school-based civic education. In N. McElvany, A. L. Grecu, R. Lorenz, M. Becker, C. Dignath, H. Gaspard & F. Lauer mann (Hrsg.), *50 Jahre Schulentwicklung – Leitthemen der empirischen Bildungsforschung* (S. 230–261). Beltz Juventa. <https://doi.org/10.17605/OSF.IO/TR8JP>

Dieser Artikel stellt keine Kopie der Veröffentlichung dar und entspricht möglicherweise nicht exakt der endgültigen und maßgeblichen Version des in der Zeitschrift veröffentlichten Artikels.

Abstract

The adoption of national educational standards in Germany in the early 2000s represented a paradigm shift away from input-driven to output-driven educational practice and research. However, no standards were adopted for civic education. Against this background, we conceptualized and operationalized a civic literacy competence model, which includes a political knowledge test as well as further dispositions in the areas of motivations, attitudes and volitions. With this study, we examine the extent to which the factorial structure of the model is supported by empirical data. In addition, measurement invariance, developmental sensitivity, and convergent and discriminant associations of the instruments were examined. Based on data from 1,047 German students attending 7th ($n = 613$, $M_{\text{age}} = 12.47$; 47.4% female; 44.4% immigrant background) or 10th grade ($n = 434$, $M_{\text{age}} = 15.49$; 47.7% female; 35.9% immigrant background), we calculated confirmatory factor analyses and mean comparisons. Furthermore, we specified and tested a nomological network. The results show that most of the theoretical assumptions are supported by the data. For political efficacy, a distinction should be made between internal and external political efficacy. In addition, the data suggest that a common factor should be assumed for the willingness to politically and the willingness to civically participate scales. The application of the civic literacy framework shall provide important insights for the research and practice of (school-based) civic education.

Keywords: Civic Education · Civic Literacy · Educational Standards · Confirmatory Factor Analysis · Measurement Invariance

Zusammenfassung

Die Verabschiedung nationaler Bildungsstandards in Deutschland in den frühen 2000er Jahren bedeutete einen Paradigmenwechsel weg von einer inputorientierten hin zu einer outputorientierten Bildungspraxis und -forschung. Für den Politikunterricht wurden jedoch keine Standards verabschiedet. Vor diesem Hintergrund wurde ein Kompetenzmodell zur politischen Bildung konzeptualisiert und operationalisiert, das neben einem Test zum politischen Wissen auch weitere Dispositionen in den Bereichen Motivationen, Einstellungen und Volitionen umfasst. In dieser Studie wird untersucht, inwieweit die faktorielle Struktur des Modells durch empirische Befunde gestützt wird. Darüber hinaus wurden Messinvarianz, Entwicklungssensitivität sowie konvergente und diskriminante Assoziationen der Instrumente untersucht. Basierend auf den Daten von 1.047 deutschen Schülerinnen und Schülern der 7. ($n = 613$, $M_{\text{age}} = 12,47$; 47,4% weiblich; 44,4% Migrationshintergrund) beziehungsweise 10. Klasse ($n = 434$, $M_{\text{age}} = 15,49$;

47,7% weiblich; 35,9% Migrationshintergrund) wurden konfirmatorische Faktorenanalysen und Mittelwertvergleiche gerechnet. Darüber hinaus wurde ein nomologisches Netzwerk spezifiziert und getestet. Die Ergebnisse zeigen, dass die meisten der theoretischen Annahmen durch die Daten gestützt werden. Beim politischen Effektivitätsgefühl sollte zwischen internem und externem politischen Effektivitätsgefühl unterschieden werden. Außerdem legen die Daten nahe, dass ein gemeinsamer Faktor für die Bereitschaft zur politischen und zivilgesellschaftlichen Partizipation angenommen werden sollte. Die Anwendung des Civic Literacy Frameworks soll wichtige Erkenntnisse für die Forschung und Praxis der (schulischen) politischen Bildung liefern.

Schlüsselwörter: Politische Bildung · Civic Literacy · Bildungsstandards · Konfirmatorische Faktorenanalyse · Messinvarianz

Introduction

In 2001, the results of the first round of the Programme for International Student Assessment (PISA) were published (OECD, 2001). In Germany, the publication of the results was accompanied by a public outrage about the poor academic performance of German students (Lange, 2002). For weeks, Germany's disappointing ranking dominated the headlines and public discussion. While the poor performance in the Trends in International Mathematics and Science Study (TIMSS) four years earlier already revealed initial deficits, the debate on the perceived qualitative decline of the German educational system and the demise of German students' international competitiveness reached new heights in the winter of 2001/02 (Waldow, 2009). In the post-‘Pisa-shock’ period, a variety of reforms were initiated, such as a massive increase in educational spending combined with better access to early childhood education, greater support for children from socioeconomically disadvantaged families, and an increase in the number of schools that offer all-day schooling (Ertl, 2006). Another reform of the post Pisa-shock period was the introduction of national educational standards (*nationale Bildungsstandards*) by the conference of ministers of education (*Kultusministerkonferenz*, KMK) for a number of school subjects (Klieme et al., 2003). Educational standards were introduced for the subjects Mathematics, German, the first foreign language, biology, chemistry and physics. Critics complained that this would ennoble these subjects as core subjects (*Kernfächer*) and declare other subjects unimportant (Köller, 2016). It was thus up to the accompanying didactic research communities of the remaining subjects to formulate educational standards.

In Germany, the Society for Civic Education Didactics and Civic Youth and Adult Education (*Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung*, GPJE) presented a proposal for national educational standards in school-based civic education (GPJE, 2004). The proposal subsequently served as the basis for further conceptualizations of educational standards and competence models in civic education in the years after the Pisa-shock, such as May's model for democratic competence (2007). However, the most well known competence model of school-based civic education in Germany is the model of political competence (*Modell der Politikkompetenz*) by Detjen et al. (2012). It is a further development of the GPJE model (2004) and the standards of political knowledge by Weißeno et al. (2010). The model is based on a relatively broad educational-psychological concept of competence (see Weinert, 2001), which includes cognitive as well as affective, motivational, and volitional dispositions. Thus, political knowledge, motivations and attitudes as well as the abilities to make

political judgments and to politically participate are dimensions of the model of political competence. The model has made a decisive contribution to the shift from normative to theory-driven and output-oriented didactics in civic education research (e.g., Massing, 2022). Despite a broad agreement within political didactics, the model has so far only been partially empirically tested. One reason for this is the operationalizability and hence the operationalization of the two dimensions ‘ability to make political judgments’ and ‘ability to politically participate’. For the few existing operationalizations, a conflation with political knowledge can be observed (see May et al., 2020). The future task of civic education research must therefore be to construct competence models that connect the normative ideas of the field with objective and empirical evidence.

While there are standardized tests for the core subjects in Germany for verifying compliance with the educational standards, empirical evidence on the effectiveness of civic education with regard to the acquisition of competence falls short of expectations (e.g., Hedinger, 2018). The civic literacy framework presented in this contribution offers a way to comprehensively and systematically consider competence in the context of civic education (Alscher et al., 2022b). It therefore aims to advance educational monitoring and evaluation in order to benefit civic education research and practice. In the framework, civic literacy is defined as (1) the knowledge and skills to make informed decisions regarding socially and politically relevant issues, and (2) the dispositions for action (i.e., motivations, attitudes and volitions) that are necessary to implement the decisions made. Consequently, the framework includes the four main areas political knowledge, motivations, attitudes, and volitions. In the future, the framework will be the basis for civics assessments in the National Educational Panel Study (*Nationales Bildungspanel*, NEPS), which aims to collect longitudinal data on competence development, educational processes, educational decisions, and educational returns in Germany.

Founded in 2009, the NEPS is a multilocal network of educational research institutions in Germany aiming to provide longitudinal data covering formal, non-formal and informal educational contexts across the whole life course (Blossfeld, Maurice & Schneider, 2019) (NEPS, 2022). Doing so, NEPS is organized in eight pillars covering the different stages of learning in the life course. Major fields of interest are the development of competencies, educational decisions, learning environments, returns to education considering the effects of social inequalities and migration backgrounds as well as educational transitions. Civic literacy is one main focus of pillar 4 which centers on lower secondary school as a critical bridge between elementary school and either the general educational or vocational higher secondary domain (Fabian et al., 2019).

The surveys conducted as part of the second version of the Start Cohort Grade 5 (SC 8) of the NEPS will enable scientists, practitioners and policy makers to follow competency trajectories longitudinally throughout lower secondary education and to gain an impression of prognostic conditional relationships.

Dimensions of Civic Literacy

Important characteristics of competencies are that they are domain-specific and learnable, and that their structure is based on the situation to which the competency is relevant (Hartig & Klieme, 2006). In German national empirical educational research, the model of political competence (Detjen et al., 2012) dominates the discourse. Internationally, the frameworks of the International Civic and Citizenship Education Study (ICCS: e.g., Schulz et al., 2022) and the framework of the Civics Assessment of the National Assessment of Educational Progress (Civic Assessment of the NAEP: e.g., National Assessment Governing Board, 2018) are of particular importance.

In ICCS 2022, political knowledge is the main aspect of the cognitive domains, which form the Civic Knowledge Framework. In addition, there are the two affective-behavioral domains Attitudes and Engagement that form the Civic Attitudes and Engagement Framework. One of the benefits of the frameworks in the ICCS is that the content is fully operationalized and tested. However, an unavoidable disadvantage of such a large international comparative study is that no reference can be made to the peculiarities of national political systems. Furthermore, assessments in ICCS 2022 are based on two different instead of one consolidated framework.

For the Civics Assessment of the NAEP, the three dimensions (1) knowledge, (2) skills, and (3) dispositions are distinguished (National Assessment Governing Board 2018). The different dimensions are assessed together across different task formats (e.g., multiple choice, short- or long-answer questions). Each of the tasks represents two to three dimensions. The results are combined to an overall score. The determination of an overall score ensures an intuitive and simple interpretation of the results. However, with only an overall score there is no information about the performance in the specific dimensions of the framework.

Using data from ICCS 2009, Hoskins et al. (2015) created a civic competence composite indicator that includes the dimensions knowledge and skills for democracy, social justice values, participatory attitudes and citizenship values. Their competence composite indicator allows calculating and evaluating an overall score as well as the performance in the specific dimensions of the indicator.

While there seems to be broad agreement among scholars that civic literacy is a multi-dimensional concept, different frameworks and models were developed leading to the investigation of different dimensions (Wray-Lake et al., 2017). One of the differences, for example, is that the abilities to make political judgments and to politically participate from the German political competence model do not seem to play a role in the international frameworks and models. This is mainly because civic literacy as such is understood as a precondition for engagement (Hoskins et al., 2015) and thus interferes with the concept of ability to politically participate and, to a lesser extent, with the ability to make political judgments. What most conceptualizations have in common, however, is political knowledge, political motivations, political attitudes, and political volitions. These are also the four dimensions of our civic literacy framework (see figure 1). Furthermore, the framework consists of facets. Two facets represent one dimension of the framework. This does not apply to the dimension of political knowledge, which has no separate facets. In contrast to educational achievement tests, the civic literacy framework is not referenced to curricula or instructional objectives (see Messick, 1984, p. 217). The framework shall provide an opportunity to measure and evaluate students' competence in the civic domain.

Political Knowledge

Political knowledge is usually considered a major tool for understanding the political world and everything related to it. It plays a vital role in the perception of the democratic system as well as current political events and is considered a desirable democratic outcome affecting voting behavior (Mondak, 2001). A widely accepted definition of political knowledge is that it encompasses “the range of factual information about politics that is stored in long-term memory” (Delli Carpini & Keeter, 1996, p. 10). As such, political knowledge is an important prerequisite for understanding political discourses and being able to classify their relevance, to process political information, and to gain an understanding about one's own political preferences (Cramer & Toff, 2017). Research finds that knowledgeable citizens engage less often in unlawful political protest activities (Gil de Zúñiga & Goyanes, 2021) and are more likely to cast their vote (Gil de Zúñiga & Diehl, 2018). However, knowledgeable citizens also have a higher tendency for opinion polarization (Herne et al., 2019) and are more likely to vote for populist parties (Stanley & Cześniak, 2022). Based on data from ICCS 2016, it was found that there are significant, positive correlations between students' political knowledge and their willingness to participate in politics in all 24 participating countries (Schulz et al., 2018). The same finding was reported for 7th and 10th grade students from North Rhine-Westphalia in Germany (Alscher et al., 2022a). A recent systematic review of research articles using the ICCS data

shows that the cognitive domain has so far received comparably little attention claiming that more in-depth analyses of political knowledge are needed (Myoung & Liou, 2022).

Political Motivations

Empirical evidence informed by the expectancy-value theory (EVT; Steinmayr et al., 2019; Wigfield & Eccles, 2000) suggests that individuals are motivated to engage in activities based on (1) the expectation of success and (2) the value they assign to the activity. In the past, EVT has proven to be a valuable resource in explaining students' effort (Dietrich et al., 2017), career choices (Gaspard et al., 2019), achievement (Lauermann et al., 2017; Meyer et al., 2019) and attitudes – for instance towards future political participation (Levy & Akiva, 2019) or attributes of a 'good' citizen (Liem & Chua, 2013). The expectation of success can be described as one's belief to accomplish something in the future she or he set out to achieve. The difference towards one's self-efficacy, namely one's belief to be *capable* to accomplish something, is very small in civic education research (Levy & Akiva, 2019). The assigned value of an activity is influenced by intrinsic value, attainment value, utility value, and cost (Eccles & Wigfield, 2002). Intrinsic value refers to the experienced enjoyment of an activity and is conceptually closely related to an individual's interest (Wigfield & Eccles, 2000). In civic education research, both political interest and the feeling of political effectiveness are important concepts and powerful predictors of desirable civic outcomes (Rasmussen & Nørgaard, 2018).

Self-efficacy can be understood as a person's conviction that he or she can cope with challenging situations (Bandura, 1977). In civic education research, political efficacy describes the extent to which an individual feels capable of influencing political debates and agenda setting, political decision-making, and political officials (Rasmussen & Nørgaard, 2018). Scholars distinguish between internal and external political efficacy. While internal political efficacy expresses one's own ability to influence the formation of political will, external political efficacy describes the perceived general responsiveness or collective suggestibility of the political system (e.g., Oberle & Wenzel, 2019). Indeed, empirical studies show that political effectiveness is related to expected future political participation (Levy & Akiva, 2019) and actual political participation (McDonnell, 2020).

Political interest is an attitude and thus analytically and empirically to be distinguished from behavior. It can be defined as the extent to which politics arouses a citizen's curiosity (van Deth, 1990). Empirical evidence suggests that political interest appears to be well approximated as a one-dimensional concept with a single underlying dimension capturing most inter-individual differences even when considering different aspects of politics (Prior, 2019). Similar to

political efficacy, political interest has proven to be an important prerequisite for expected future political participation (Alscher et al., 2022a) and actual political participation (Prior & Bougher, 2018; Stattin & Amnå, 2022).

Political Attitudes

Attitudes are mental tendencies expressed by valuing a certain object with a certain degree of affection or rejection (Eagly & Chaiken, 1993). Political attitudes can be understood as acquired, enduring, often more strongly emotional orientations toward classes of persons or objects with a connection to political issues (Detjen et al., 2012). Attitudes toward (state-relevant) objects can be represented by an individual's opinion about democracy while attitudes toward (social groups or classes of) people can be observed through an individual's opinion regarding pluralism (Alscher et al., 2022b). For the opinion about democracy, a distinction must be made between the general agreement with the idea of democracy and its constituent principles on the one hand and the evaluation of the current state of democracy in a particular country and time period on the other hand (Gabriel, 2000).

According to ICCS 2009 and 2016 data, adolescents from countries with more inclusive and multiculturalist policies showed more positive attitudes towards ethnic minorities than adolescents from countries with less inclusive and multiculturalist policies (Kim & Byun, 2019). With the exception of a shift towards more inclusive attitudes in terms of racial and gender equality, political attitudes remained rather similar between the 1999 and 2009 CIVED/ICCS cohorts (Barber & Ross, 2019). Furthermore, ICCS 2009 data shows that the introduction of inequality on the basis of achievement through academic segregation impedes positive attitudes towards social cohesion and other democratic values (Kavadias et al., 2017)

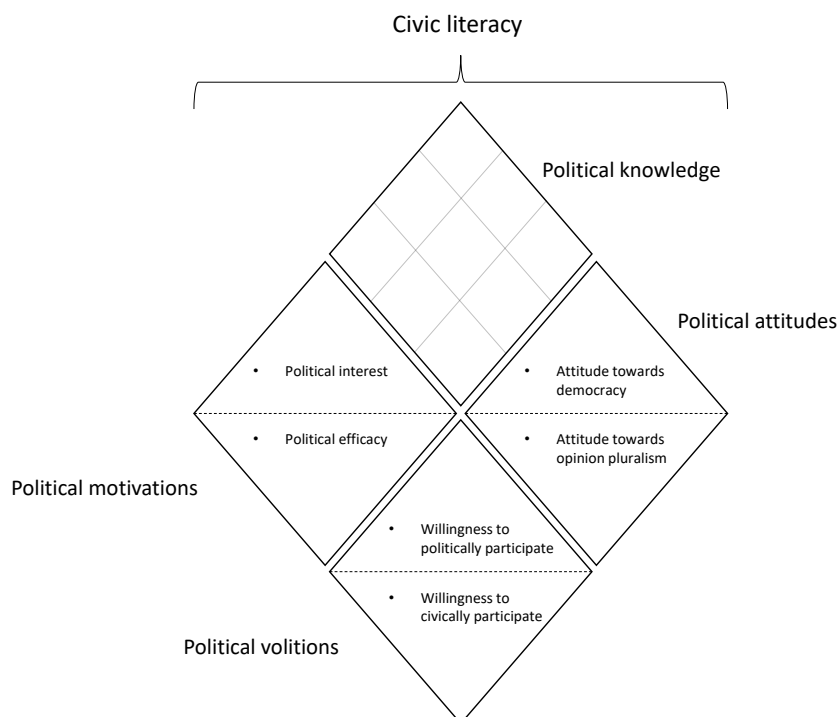
Political volitions

Volitions are cognitive processes that are related to the concrete realization of action goals in the context of the preactional and actional phases of action (Achtziger & Gollwitzer, 2010). In general, intentions are considered immediate antecedents of actual behavior (Ajzen, 1991). In civic education research, too, political volitions are viewed as the closest proxy indicator for actual political behavior (Quintelier & Hooghe, 2013). Political volitions represent an individual's willingness to participate in political or civic life including activities such as voting, protesting or joining a political party. Looking at youth's intentions rather than their actual participation is the only way to get a better understanding of minors' future participation. This especially includes activities that are either not available to them (e.g., voting), for which they need support (e.g., money from parents to be able to donate) or which are very unusual at a

young age (e.g., joining a political party) (see McWhirter & McWhirter, 2016). From a theoretical point of view, participation activities can be divided into more civically oriented or latent political participation, and manifest political participation (Ekman & Amnå, 2012). Empirical evidence from the U.S and Italy indeed suggests a strong relationship between voting intentions and actual voting behavior (Ajzen et al., 1982; Roccato & Zogmaister, 2010). Furthermore, CIVED 1999 data suggests that school council involvement and volunteering are associated with greater expectations to become an informed voter and an active citizen (Homana, 2018).

Figure 1

Civic Literacy Framework



Note. From „Civic literacy — about the theory and measurability of competence in school-based civic education” by Alscher et al., 2022, *Zeitschrift für Erziehungswissenschaft* (<https://doi.org/10.1007/s11618-02201085-0>). CC BY 4.0.

Current Study

The development of a competence model that enables the extensive assessment and evaluation of students' civic literacy shall make an important contribution to the optimization of civic education and its teaching practices. So far, there are hardly any fully operationalized civic literacy

models. In addition, systematic analyses of the dimensions and facets of civic literacy are missing. With this study, we aimed to evaluate the degree to which our data supports the structure and assumptions of the proposed framework (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014) and therefore investigated the following research questions:

1. Is the theoretically assumed structure of the framework better supported by the data than alternative, simpler models?

1.1 Are the seven facets modeled in the framework empirically separable into distinct factors?

Given that the various facets of the framework represent complex and theoretically separable concepts, we expected that all facets are separable, but interrelationships (e.g. common factors) exist. In addition, we expected that all items are at least congeneric items, i.e., the factor loadings and error variances of the items may be different, but all items' factor loadings reach substantial magnitude ($\lambda \geq .50$).

1.2. Is a hierarchical model with facets and dimensions better supported by the data than a non-hierarchical model?

We assumed that a hierarchical model is better supported by the data than a non-hierarchical model. The reason for this is the distinction between broader dimensions and more concrete facets.

2. What levels of measurement invariance are present for the framework for 7th and 10th grade?

We expected that configural, metric and scalar invariance holds indicating that the scores from each group have the same unit of measurement (factor loadings) and the same origin (intercepts). Hence, the latent means could be compared across 7th and 10th grade.

3. What differences can be found for the mean scores of the framework's facets between 7th and 10th grade?

We assumed that all facets are developmentally sensitive and thus more pronounced (i.e., have higher mean values) at the 10th grade level than at the 7th grade level. In addition, we assumed that political knowledge will be particularly higher, because knowledge transfer is a central goal and recurring component of (civic) education.

4. How closely are the framework's facets related to other facets of the framework (providing convergent evidence) and constructs that are unrelated to civic literacy (providing discriminant evidence)?

We expected that the different facets of the framework are more closely related to each other than to sense of social affiliation, which is theoretically unrelated to civic literacy. Since the framework distinguishes between facets and dimensions, we also assumed that facets attributed to the same dimension are more strongly related to each other than facets belonging to different dimensions.

Methods

Participants and Procedure

Participants were 1,047 students attending 7th ($n = 613$, $M_{\text{age}} = 12.47$; 47.4% female; 44.4% immigrant background) or 10th grade ($n = 434$, $M_{\text{age}} = 15.49$; 47.7% female; 35.9% immigrant background). Data were collected in fall and winter 2021/2022 in the context of the Study on the Development of Political and Civic Competence in Adolescence (*Studie zur Entwicklung politischer und gesellschaftlicher Kompetenz im Jugendalter*, EPKO). The schools were located in rural, semi-urban and urban areas of West-Germany in the federal state of North Rhine-Westphalia. The response rate was 72%. Students not reached either did not have informed consent or were absent on the test day due to illness or quarantine measures. The share of students with an immigrant background was with 41.8% comparable to the official percentage of students with an immigration background in the same federal state in 2020 (40.0%, MSW, 2022).

The assessments took place in the morning and took about two and a half hours including breaks. At the beginning of the assessments, the students worked on the knowledge test for one hour. After that, students took an 18-minute matrix test and a six-minute reading test to validate the knowledge test. Finally, all participants completed a questionnaire for about 40 minutes. The study was conducted in compliance with the German Research Foundation's (*Deutsche Forschungsgemeinschaft*, DFG) guidelines for good scientific practice. In addition, the university's Joint Ethics Committee examined and approved the research project. Students could only participate in the study if they presented an informed consent signed by their parents at the beginning of the assessment. Without informed consent, students had to leave the classroom. Each class was awarded 50 EUR for their participation.

Measures

Political Knowledge

The political knowledge test instrument included 63 items at the 7th grade level and 70 items at the 10th grade level. 34 items were answered by students in both grades, so that a total of 99 items were used. Some of the test items were taken from the POWIS-study (Goll et al., 2011) and adjusted if needed. The instrument targets different facets and content areas of political knowledge (Alscher et al., 2022b). In order to link the test results of the 7th and 10th grade, we calculated a unidimensional multigroup 1-parameter model using R and the “TAM” package (Robitzsch et al., 2021). Before calibration, we detected rapid guessing behavior, which led to the deletion of 6.76% of the answers. We set the threshold for deletion at seven seconds of processing time. The threshold represents the local minimum between the two modes in the response time distribution (Wise, 2017). For every participant, we estimated a person parameter (θ). Overall and grade-specific reliabilities are shown in table 1.

Political Motivations

Political Interest. We measured political interest with the Short Scale Measuring Political Interest (*Politisches Interesse Kurzskala*, PIKS, Otto & Bacherle, 2011). Due to space constraints and after psychometric consideration of a prior assessment (Alscher et al., 2022a), only the first four items of the scale were used (e.g., “I observe political events with great interest.”). The introductory question in all motivation and attitude scales read as follows: “To what extent do the following statements apply to you?” The students were asked to rate each item on a four point scale ranging from 1 = “doesn’t apply at all” to 4 = “applies very well”.

Political Efficacy. We measured political efficacy with the Political Efficacy Kurzskala (PEKS, Beierlein et al., 2014). The introductory question read as follows: “To what extent do the following statements apply to you?” The scale consists of two items on internal (e.g., “I am good at understanding and assessing important political issues.”) and two items on external (e.g., “Politicians strive to keep in close touch with the people.”) political efficacy.

Political Attitudes

Attitude towards Democracy. We measured attitude towards democracy with a newly constructed four-item-scale. The introductory question read as follows: “When you compare democracy with other forms of government, how much do you agree with the following statements?” The items of the scale are (1) “Democracy is the best form of government there is.”, (2) “Democracy best ensures a just and fair society.”, (3) “Democracy is the best way to govern a country.”, and (4) “Democracy best protects fundamental freedoms.”

Attitude towards opinion pluralism. We also measured attitude towards opinion pluralism with a novel four-item-scale. The introductory question read as follows: “To what extent do the following statements apply to you?” The items of the scale are (1) “All people should have the same opportunities to stand up for their opinions.”, (2) “The opinion of minorities should also be taken into account in decision-making.”, (3) “Dissenting opinions should be accepted and respected.”, and (4) “Everyone should have the right to stand up for their opinion, even if the majority disagrees.”

Political Volitions

Willingness to politically participate. We measured the willingness to politically participate with a four-item-scale. The introductory question read “If you want to voice your opinion, how likely are you to take the following actions?” For each item, students chose a response on a four-point scale comprising 0 = “not at all likely”, 1 = “rather not likely”, 2 = “rather likely”, and 3 = “very likely”. Each item refers to a different political activity, such as voting, protesting, signing a petition.

Willingness to civically participate. We also measured the willingness to civically participate with a four-item-scale. The introductory question also read “If you want to voice your opinion, how likely are you to take the following actions?” Each item refers to a different civic activity, such as donating money, volunteering, boycotting products.

Control Variables

Sense of social affiliation. We measured the sense of social affiliation with a scale from the NEPS (Bömmel et al., 2022) including five items (e.g., I find it easy to make new friends”). The introductory prompt read “For each statement, please tell me how it applies to you using the scale.” The students were asked to rate each item on a four point scale ranging from 1 = “doesn’t apply at all” to 6 = “applies very well”.

School grade. 7th grade was coded as 0 and 10th grade as 1.

Data Analysis

We conducted all analyses in R, mostly using the “lavaan” and “psych” packages. Confirmatory factor analyses were performed to investigate the internal structure of the model. For this purpose, we compared a series of nested models with each other. The first three models were non-hierarchical models and contained either facets (1), dimensions (2), or a global factor (3). Additionally, three hierarchical models contained either the facets (4) or the dimensions (5) and a global factor, or facets and dimensions but no global factor (6). We set the variance of the

factors to one. All models included all items. We disallowed correlations between items in all models. Correlations between factors were allowed only when no other, higher factor was modelled. Prior to the CFAs, we tested the fitness of the dataset for performing factor analyses using the Kaiser-Meyer-Olkin test and the Bartlett's Test of Sphericity. Based on the CFAs' results, the model was adjusted. We used the adjusted model for further analyses.

For the examination of measurement invariance between students in 7th and 10th grade, we performed a series of multiple group confirmatory factor analyses (MGCFA), which tests the invariance of estimated parameters of two nested models across different groups (Cheung & Rensvold, 2002). The first MGCFA was specified for each group separately without any equality constraints. For this baseline model, we assessed the overall model fit to examine whether configural invariance holds. In a second model, we constrained factor loadings to be equal across 7th and 10th graders to check for metric invariance. By comparing the fit of the baseline model and the metric model, we can determine whether metric invariance holds. After establishing metric invariance, scalar invariance can be tested by specifying a third model with both constrained factors and intercepts. By comparing the metric model and the scalar model, we can determine whether scalar invariance holds. A common statistic for assessing differences in fit between two nested models is the Likelihood Ratio Test. However, χ^2 difference tests were criticized for their dependence on sample size (e.g., Lee & Smith, 2020). Hence, we also used differences in the comparative fit index (Δ CFI) to evaluate meaningful change in model fit (Cheung & Rensvold, 2002). When sample size is rather large ($N > 300$) and sample sizes are comparable across groups, changes in the root-mean-square error of approximation (RMSEA), and the standardized root mean square residual (SRMR) can also be evaluated to identify noninvariance (Chen, 2007). For testing metric invariance, a change of $\geq -.010$ in CFI, supplemented by a change of $\geq .015$ in RMSEA or a change of $\geq .030$ in SRMR would indicate noninvariance. For the examination of scalar invariance a change of $\geq -.010$ in CFI, supplemented by a change of $\geq .015$ in RMSEA or a change of $\geq .010$ in SRMR would indicate noninvariance (Chen, 2007).

In contrast to achievement tests, changes in competency measures are not only attributed to instructional sensitivity (Ing, 2018). In competence research, both school and non-school characteristics are considered as possible sources of development (Messick, 1984). Together, these characteristics represent the developmental sensitivity of a test. We tested the developmental sensitivity of the framework by comparing the mean values between 7th and 10th graders using an unpaired two-samples t-test. Before comparison, we weighted students' responses based on type of school.

We obtained convergent and discriminant evidence by specifying and evaluating a nomological network (Cronbach & Meehl, 1955). For this purpose, we identified three nomological orders. The first order represented correlations between facets from the same dimension of the framework. The second order represented correlations between facets from different dimensions of the framework. To this end, the facets of the framework were correlated with political interest, which is a central facet of the framework. Exceptions are political effectiveness and political interest itself. These two facets were correlated with attitude towards democracy. The third order represented correlations between facets of the framework and sense of social affiliation, which is theoretically unrelated to civic literacy. We compared the magnitude of correlations of the three orders with each other. Any differences in correlations were tested for statistical significance.

In this study, we evaluated global model fit with the comparative fit index (CFI), the Tucker-Lewis Index (TLI), the RMSEA, and the SRMR, using the following fit criteria for acceptable fit: $CFI \geq .900$, $TLI \geq .900$, $RMSEA \leq .100$, $SRMR < .100$ (Weston & Gore, 2006). In addition, χ^2 , the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used for comparative model fit. In all models, we treated political knowledge as a latent variable with a single indicator. We controlled for the clustered structure of the data by using class ($n = 69$) as the cluster variable, thus enabling the computation of robust standard errors of the parameters estimated in the CFAs. We handled missing data through full information maximum likelihood (FIML) method. Supplemental material (e.g. the full scales, additional analyses) and the script are available at <https://osf.io/tr8jp/>.

Results

Descriptive Statistics

Table 1 shows the latent correlation coefficients, mean, standard deviations and range as well as the response rate and reliability for each of the construct. The mean score of the political knowledge is slightly above 0. This is because in the process of scaling the average score of the 7th graders was set to 0. The average score of the 10th graders is 0.53. Average scores for the other variables are a little above the midpoint of the response scales. The only exception is the attitude towards opinion pluralism, whose average value of 3.39 is well above the midpoint and close to the maximum value of 4.00.

All 1,047 participants completed the political knowledge test. The reason for the quantitative variance in the response rates of the different constructs is the positioning of the

instruments in the questionnaire. Due to internal school processes and stricter hygiene standards imposed by the pandemic, it was not possible to ensure that all participants were able to complete all questions in every survey for time reasons.

As expected, the bivariate correlations in Table 1 show moderate to strong, statistically significant correlations. The correlations between political interest and political efficacy as well as the correlations between willingness to politically and willingness to civically participate are particularly high. Again, this is not surprising if considering that in both cases both facets belong to the same dimension (i.e., political motivations and political volitions).

The WLE reliability for the knowledge test was good. The McDonald's ω for almost all other scales were acceptable to excellent ($\omega \geq .70$). The exceptions were attitude towards opinion pluralism across the two grades ($\omega = .66$) and in the 10th grade ($\omega = .63$), and willingness to civically participate in the 10th grade ($\omega = .68$).

Table 1

Means, standard deviations, and correlations

Variable	1	2	3	4	5	6	7	8
1. Political knowledge	–							
2. Political interest ^a	.36	–						
3. Political efficacy ^a	.42	.81	–					
4. Attitude towards democracy ^a	.57	.38	.50	–				
5. Attitude towards opinion pluralism ^a	.39	.18	.25	.39	–			
6. Willingness to politically participate ^a	.36	.42	.46	.31	.30	–		
7. Willingness to civically participate ^a	.36	.30	.31	.27	.31	.98	–	
8. Sense of social affiliation	.01	.14	.21	.14	.21	.17	.18	–
<i>M</i>	0.23	2.17	2.36	2.95	3.39	2.38	2.49	2.83
<i>SD</i>	0.86	0.76	0.59	0.75	0.68	0.71	0.66	0.55
Actual Range	^b 1.7 – 3.1	1 – 4	1 – 4	1 – 4	1 – 4	1 – 4	1 – 4	1 – 4
<i>n</i>	1047	994	948	947	867	750	753	750
McDonald's ω	^c .89	.88	.78	.89	.66	.78	.72	.72
Grade 7	^c .89	.88	.79	.89	.91	.79	.76	.71
Grade 10	^c .88	.89	.76	.89	.63	.76	.68	.75

Note. *M* and *SD* represent mean and standard deviation, respectively.

^a*M* and *SD* are calculated using the row means of the variable indicators. ^brounded to one decimal place. ^cInstead of McDonald's ω , the WLE reliability is reported.

Correlations printed in bold reached statistical significance at the 5 % level.

Framework Evaluation

We tested the factorability of the data with the Kaiser-Meyer-Olkin test (KMO). The KMO value for the present data is .90 and exceeds the suggested cutoff value of .60, indicating that a factor analysis can probably be conducted (Kaiser & Rice, 1974). The sphericity of the data was tested with Bartlett's Test of Sphericity. The test suggests that there is probably sufficient significant correlation in the data in order to perform a factor analysis ($\chi^2(300) = 7768, p < .001$).

For the comparison of the models, model 1 was used as the baseline model. Table 2 shows that model 1 and model 6 outperform all other models with regard to different goodness of fit indicators. The χ^2 test suggests that model 1 fits the data the best among the specified models. The AIC also favors model 1. However, the BIC favors model 6 over model 1 ($\Delta\text{BIC} = -22$). In contrast to the AIC, the BIC awards a bigger penalty for additional parameters and thus favors parametrically simpler models. Model 6 has fewer correlations and thus fewer parameters than model 1.

Table 2

Comparison of Fit Indices in different nested CFAs

Model	χ^2			Robust RMSEA		AIC	BIC	Δ AIC	Δ BIC
	Value	df	p^a	Value	90% CI				
M1: Facets only	989	256	–	.051	[.047, .055]	49121	49587	–	–
M2: Dimensions only	2799	271	< .001	.094	[.091, .097]	50901	51292	1780	1705
M3: G-factor only	5370	275	< .001	.134	[.131, .137]	53565	53936	4444	4349
M4: Facets & g-factor	1706	270	< .001	.070	[.067, .074]	49810	50206	689	619
M5: Dimensions & g-factor	3017	273	< .001	.097	[.094, .101]	51116	51497	1995	1910
M6: Dimensions & facets	1029	265	< .001	.051	[.048, .055]	49144	49565	23	-22

Note. RMSEA = root-mean-square error of approximation; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.

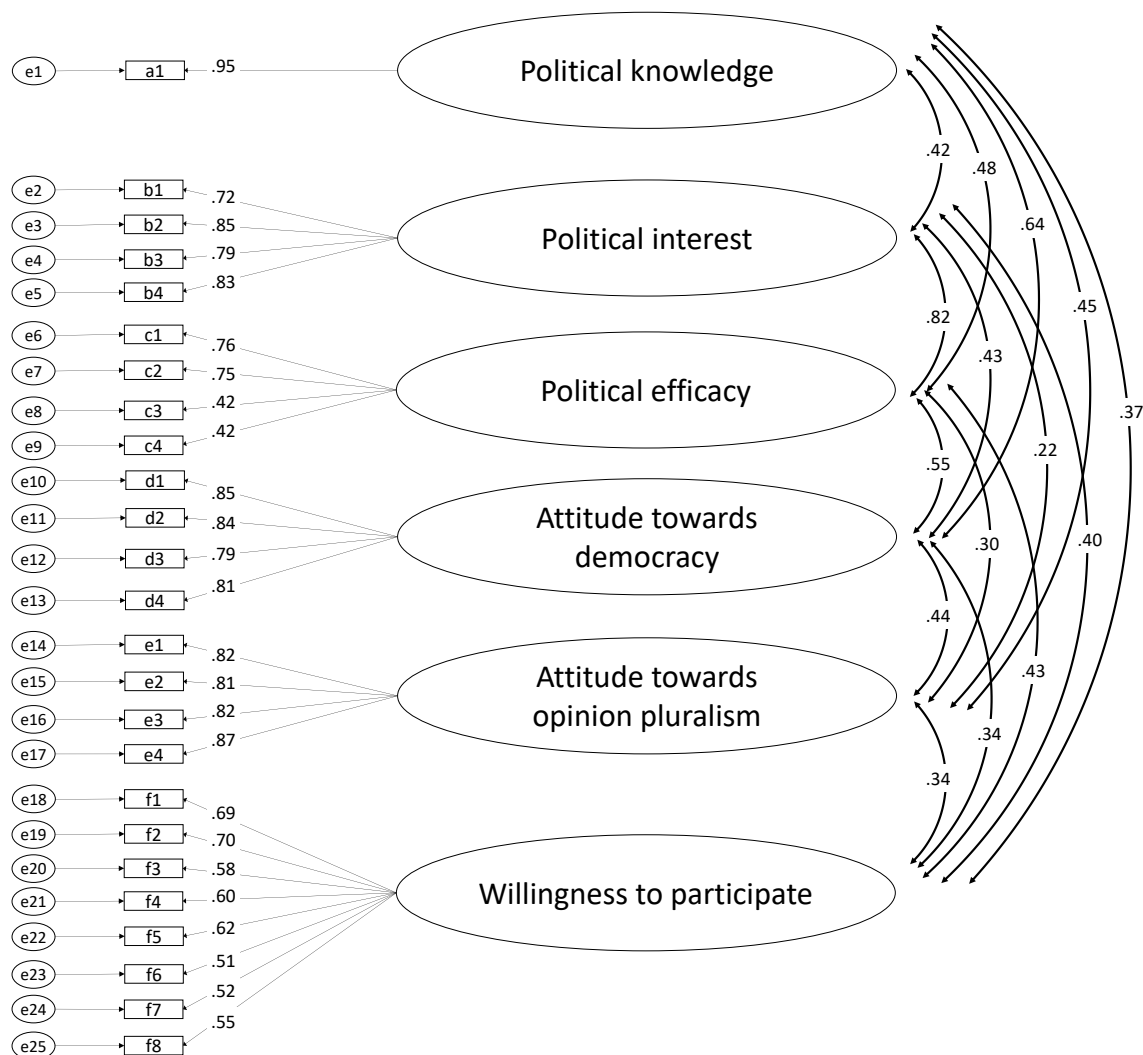
^aThe χ^2 of Model 1 is used as the benchmark.

Model 1 has an overall acceptable global model fit (CFI: .926; TLI: .913; RMSEA: .052; SRMR: .077; see figure B1). All factor loadings reached significance at the 5% level. In addition, almost all factor loadings reached substantial magnitude ($\lambda \geq .50$) except for items c3 and c4. While items c1 and c2 loaded highly on political efficacy, items c3 and c4 showed rather low loadings on political efficacy ($\lambda = .41$ and $.41$ respectively). Eventually, factor loadings for the items of the willingness to civically participate facet were overall only moderate in size (.50

$< \lambda < .61$). Furthermore, the correlation between the two willingness facets is particularly high suggesting that the two factors might measure the same construct. We found similar patterns of results in model 6 (see figure B2). There too, the items of the willingness to civically participate have rather moderate factor loadings. Furthermore, items c1 and c2 again load more strongly on the political efficacy facet than items c3 and c4. Eventually, the modification indices added to the impression, that the political efficacy scale and the two willingness scales cause problems.

Figure 2

Confirmatory factor analysis (adjusted model)



Note. CFI: .925; TLI: .914; RMSEA: .052; SRMR: .078.

All factor loadings and correlations reached statistical significance at the 5% level.

Based on these findings, we adjusted model 1. We combined the two willingness facets into one single facet. However, we were unable to separate the political efficacy facet into an internal and an external political efficacy facet, because there are not enough items available in

the current framework. The adjusted model (see figure 2) was compared to model 1 and model 6 (see table B1). The χ^2 and AIC still favor model 1, but the BIC now favors the adjusted model. All further analyses will be carried out using the adjusted model.

For the testing of measurement invariance by class grade, we used the adjusted model as the baseline model. The configural model exhibited an acceptable model fit (see table 2). Regarding metric invariance, the χ^2 test suggests that the configural model does not fit the data statistically significantly better than the metric model. In addition, neither the drop in the CFI ($\Delta\text{CFI} = .000$) or the RMSEA ($\Delta\text{RMSEA} = .000$), nor the drop in the SRMR ($\Delta\text{SRMR} = -.010$) exceed the suggested threshold indicating metric invariance. In terms of scalar invariance, the χ^2 test suggests that the metric model fits the data statistically significantly better than the scalar model. However, the drop in the CFI is below the .007 threshold ($\Delta\text{CFI} = .005$) and the CFI itself remains on an acceptable level. Furthermore, the RMSEA ($\Delta\text{RMSEA} = .000$) and the SRMR ($\Delta\text{SRMR} = .000$) remain stable.

Table 3

Measurement invariance between 7th and 10th grade

Model	χ^2			CFI	RMSEA	SRMR
	Value	df	p^a			
1. Configural	1338	524		.918	.054	.095
2. Metric model	1357	542	.096	.918	.054	.085
3. Scalar model	1425	561	< .001	.913	.054	.085

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root mean square residual.

^aThe χ^2 of the metric model was compared to the χ^2 of the configural model. The χ^2 of the scalar model was compared to the χ^2 of the metric model.

Next, we tested the developmental sensitivity of the framework by comparing the facets' means between the 7th and 10th grade. All means were higher in the 10th grade than in the 7th grade (see table A1). The differences between the means were statistically significant at the 5% level for political knowledge, political interest, attitude towards democracy and attitude towards opinion pluralism, and not significant for political efficacy and willingness to participate.

The correlations between the facets of the framework and sense of social affiliation largely correspond to the assumed nomological network (see table A2). The correlational relations of the first order are the largest for all facets while the relations of the third order are the

smallest. The only exception is attitude towards opinion pluralism, for which the third order relation is larger than the second order relation. All other differences in correlations were significant at the 5% level with two exceptions. The correlation between attitudes towards democracy and attitude towards opinion pluralism ($r = .39, p < .001$) was statistically only insignificantly higher ($\Delta = .01, p = .850$) than the correlation between attitude towards democracy and political interest ($r = .38, p < .001$). Moreover, the correlation between willingness to participate and political interest ($r = .37, p < .001$) was statistically only insignificantly higher ($\Delta = .20, p = .920$) than the correlation between willingness to civically participate and sense of social affiliation ($r = .17, p = .008$).

Discussion

Gradual declines of democratic regime attributes characterize contemporary autocratization (Lührmann & Lindberg, 2019). It is therefore, perhaps more important than ever, to understand what constitutes a civically literate citizen and to unravel what factors play a role in the development of young people on their way to becoming democratic and active citizens (Oberle, 2022). However, previous research has usually focused on separate areas of civic literacy, thereby often relegating either cognitive or affective dimensions to a different sphere. Thus, holistic overviews and conceptualizations of youth's civic literacy are uncommon (see Keegan, 2021). In addition, existing conceptualizations are usually rooted solely in theoretical considerations, while empirical examinations of the proposed frameworks are absent. Before data collection, a panel of experts consisting of six professors for political didactics as well as the study director of an international large-scale assessment on civic education assessed the adequacy of the civic literacy framework's content (Alscher et al., 2022b).

The results of this study show that the structure of the suggested civic literacy framework fits the data overall well and most expectations are met. Regarding the internal structure of the model, correlations, and the CFAs give rise to the assumption that internal and external political efficacy should be viewed as two separate constructs. Indeed political efficacy is often understood as a bidimensional belief that includes distinct external and internal components (Ardèvol-Abreu et al., 2020; Oberle, 2018). Previous research has shown that internal and external efficacy have different meanings with regard to the political socialization of young people. For instance, the education effect on internal efficacy is considerably more confounded by personality traits than the educational effect on external efficacy (Rasmussen & Nørgaard, 2018). Future studies should make sure to assess internal and external political efficacy separately and with enough items. Furthermore, the willingness to politically and civically

participate correlate highly and both load particularly high on the political volitions dimension. In addition, the item loadings are only mediocre. We therefore adjusted the framework accordingly (see figure 3).

Using the adjusted model, scalar measurement invariance was established according to the ΔCFI , $\Delta RMSEA$ and $\Delta SRMR$. Furthermore, the overall model fit remains rather high across the different MGCFAs.

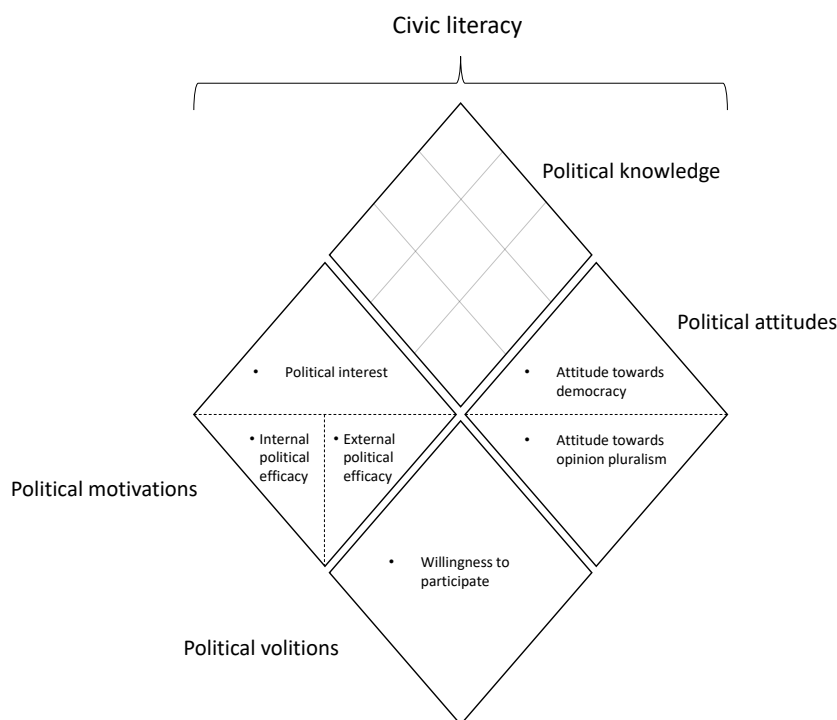
Unlike previous studies, the framework is designed to enable a longitudinal assessment of civic literacy. For this purpose, it is important that the framework is developmentally sensitive and reflects processes of students' political socialization. The comparison of the 7th and 10th grade cohort shows that all framework facets' means are higher in the 10th grade than in 7th grade. However, magnitudes of the differences vary with regard to statistical significance and effect sizes. Especially political knowledge seems to increase from 7th to 10th grade. This is not surprising when one considers that it is primarily cognitive processes and thus knowledge transfer that play a central role in civic education, whereas aspects of political practice are at best of marginal importance (Widmaier & Nonnenmacher, 2011). We can observe a mediocre difference for political attitudes. However, one of the limitations here is that the identified difference might be confounded by the fact that 7th graders are less familiar with the terms that were used (e.g., democracy, minority/majority) and therefore might have a stronger tendency for the midpoint of the scale. Rather counterintuitively, we found only small differences for political efficacy and willingness to participate. Possible reasons for this could be that younger students might have lower estimates of potential costs for meaningful engagement, assign greater value to engagement or expect higher rewards from actively participating in civic and political life. Moreover, students' decision to become active might not (only) be associated with their absolute but instead their relative level of political knowledge, attitudes or abilities compared to their peers. Hence, participation patterns could be different within a particular class and grade level but rather equal across classes and grade levels. The general mechanism is also known as the 'big fish little pond'-effect, which has already proven helpful in explaining differences in inequality of civic and political engagement (Witschge, 2022). Future research should strive for a better understanding of this relationship.

The investigation of the nomological network suggests that facets of the framework are more related to each other than to a construct that is theoretically unrelated to civic literacy. Furthermore, facets of the same dimension are usually more strongly associated with each other than facets of different dimensions. The framework's constructs thus demonstrate convergence and discrimination. The only exception is attitude towards democracy, which has very similar

correlations with attitude towards opinion pluralism and with political interest. Upon closer examination, the bigger problem seems to be the convergence between the two attitude constructs, since the correlation between the two is comparatively weak. This impression is also confirmed by the CFA of model 6 (dimensions and facets), which suggests that the two attitude facets load rather differently on the political motivation dimension. A possible reason for this could be that attitude towards opinion pluralism is less strongly concerned with original political topics. Instead, a social aspect resonates in this facet since the exchange of opinions is something that students experience regularly outside the political arena in their everyday life. This needs to be taken into account in future applications of the framework but also when other political motivations are assessed.

Figure 3

Adjusted Civic Literacy Framework



Note. Adapted from „Civic literacy — about the theory and measurability of competence in school-based civic education” by Alscher et al., 2022, *Zeitschrift für Erziehungswissenschaft* (<https://doi.org/10.1007/s11618-02201085-0>). CC BY 4.0.

Conclusion

While the empirical assessment of competencies in the school context has gained in importance in the past two decades, there are comparatively few studies with the aim of systematically

assessing civic literacy. We therefore aimed to realign the conceptual and empirical foundations of civic literacy research by empirically examining the conceptualized civic literacy framework. The results show that the framework fits the data overall well and most theoretical assumptions are confirmed. However, political efficacy should be distinguished between external and internal political efficacy. In contrast, there is no clear differences between the willingness to politically and the willingness to civically participate. The framework outlined in this study is to be applied in future studies of political socialization as well as empirical educational research and political psychology. This includes, in particular, the National Educational Panel Study (NEPS; Blossfeld and Roßbach 2019), for which civic literacy was included in the canon of competencies to be assessed in grades 7 and 10. The application of the civic literacy framework, or parts of it, shall provide important insights for future research and practice of (school-based) civic education.

References

- Achtziger, A., & Gollwitzer, P. M. (2010). Motivation und Volition im Handlungsverlauf. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln* (4th ed., pp. 309–335). Springer. https://doi.org/10.1007/978-3-642-12693-2_12
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I., Timko, C., & White, J. (1982). Self-monitoring and the attitude-behavior relation. *Journal of Personality and Social Psychology*, *42*, 426–435.
- Alscher, P., Ludewig, U., & McElvany, N. (2022a). Civic education, teaching quality and students' willingness to participate in political and civic life: Political interest and knowledge as mediators. *Journal of Youth and Adolescence*, *51*(10), 1886–1900. <https://doi.org/10.1007/s10964-022-01639-9>
- Alscher, P., Ludewig, U., & McElvany, N. (2022b). Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung. *Zeitschrift Für Erziehungswissenschaft*, *25*(5), 1221–1241. <https://doi.org/10.1007/s11618-022-01085-0>
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *The standards for educational and psychological testing*. https://www.testingstandards.net/uploads/7/6/6/4/76643089/-standards_2014edition.pdf
- Ardèvol-Abreu, A., Gil de Zuniga, H., & Gámez, E. (2020). The influence of conspiracy beliefs on conventional and unconventional forms of political participation: The mediating role of political efficacy. *British Journal of Social Psychology*, *59*(2), 549–569.
- Barber, C. & Ross, J. (2018). Cross-Cohort Changes in Adolescents' Civic Attitudes from 1999 to 2009: An Analysis of Sixteen Countries. *Child Indicators Research*, *11*, 681–703. <https://doi.org/10.1007/s12187-017-9452-0>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)
- Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2014). *Political Efficacy Kurzsкала (PEKS)*. Zusammenstellung sozialwissenschaftlicher Items und Skalen. <https://doi.org/10.6102/ZIS34>
- Blossfeld, H.-P., Maurice, J. von, & Schneider, T. (2019). The national educational panel study: need, main features, and research potential. In H.-P. Blossfeld & H.-G. Roßbach (Eds.),

- Edition ZfE: Vol. 3. Education as a Lifelong Process: The German National Educational Panel Study (NEPS)* (pp. 1–16). Springer.
- Bömmel, N., Kroh, J., Gebel M., & Heineck G. (2022). *Social Inclusion as Return to Education in the National Educational Panel Study (NEPS): Conceptual Framework and Measurement* [Manuscript submitted for publication]. National Educational Panel Study.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheung, G. W. & Rensvold, R. B. (2002) Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance, *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255, https://doi.org/10.1207/S15328007SEM0902_5
- Cohen, J. (1992). Quantitative Methods in Psychology: A Power Primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cramer, K. J., & Toff, B. (2017). The fact of experience: Rethinking political knowledge and civic competence. *Perspectives on Politics*, 15(3), 754–770. <https://doi.org/10.1017/S1537592717000949>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Detjen, J., Massing, P., Richter, D., & Weißeno, G. (2012). *Politikkompetenz – ein Modell*. Springer.
- Dietrich, J., Viljaranta, J., Moeller, J., & Kracke, B. (2017). Situational expectancies and task values: Associations with students' effort. *Learning and Instruction*, 47, 53–64. <https://doi.org/10.1016/j.learninstruc.2016.10.009>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes* (2.th ed.). Harcourt Brace Jovanovich.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.
- Ekman, J., & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3), 283–300. <https://doi.org/10.2478/s13374-012-0024-1>
- Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32(5), 619–634.

- Fabian, P., Goy, M., Jarsinski, S., Naujokat, K., Prosch, A., Strietholt, R., Blatt, I., & Bos, W. (2019). Transition and development from lower secondary to upper secondary school. In H.-P. Blossfeld & H.-G. Roßbach (Eds.), *Edition ZfE: Vol. 3. Education as a Lifelong Process: The German National Educational Panel Study (NEPS)* (pp. 231–252). Springer.
- Gabriel, O. W. (2000). Demokratische Einstellungen in einem Land ohne demokratische Traditionen? Die Unterstützung der Demokratie in den neuen Bundesländern im Ost-West-Vergleich. In J. W. Falter, O. W. Gabriel, & H. Rattinger (Eds.), *Wirklich ein Volk? Die politischen Orientierungen von Ost- und Westdeutschen im Vergleich* (pp. 41–77). Springer. https://doi.org/10.1007/978-3-663-05704-8_2
- Gaspard, H., Wille, E., Wormington, S. V., & Hulleman, C. S. (2019). How are upper secondary school students' expectancy-value profiles associated with achievement and university STEM major? A cross-domain comparison. *Contemporary Educational Psychology*, 58, 149–162. <https://doi.org/10.1016/j.cedpsych.2019.02.005>
- Gil de Zúñiga, H., & Diehl, T. (2018). News finds me perception and democracy: Effects on political knowledge, political interest, and voting. *New Media & Society*, 21(6), 1253–1271. <https://doi.org/10.1177/1461444818817548>
- Gil de Zúñiga, H., & Goyanes, M. (2021). Fueling civil disobedience in democracy: WhatsApp news use, political knowledge, and illegal political protest. *New Media & Society*, 14614448211047850. <https://doi.org/10.1177/14614448211047850>
- GPJE. (2004). *Anforderungen an Nationale Bildungsstandards für den Fachunterricht in der Politischen Bildung an Schulen: Ein Entwurf*. Wochenschau Verlag. <http://gpje.de/wp-content/uploads/2017/01/Bildungsstandards-1.pdf>
- Goll, T., Richter, D., & Weißeno, G. (2011). Politisches Wissen von Schüler/-innen mit und ohne Migrationshintergrund (POWIS): Ergebnisse einer Studie. In D. Lange (Eds.), *Schriftenreihe der DVPB, Wochenschau Wissenschaft, Entgrenzungen: Gesellschaftlicher Wandel und politische Bildung* (pp. 126–131). Wochenschau Verlag.
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Ed.), *Leistung und Leistungsdiagnostik* (pp. 127–143). Springer.
- Hedinger, F. (2018). Erkenntnisse aus der empirischen Politikwissenschaft für das Politikkompetenzmodell. In B. Ziegler & M. Waldis (Eds.), *Politische Bildung in der Demokratie: Interdisziplinäre Perspektiven* (pp. 57–73). Springer. https://doi.org/10.1007/978-3-658-18933-4_5

- Herne, K., Christensen, H. S., & Grönlund, K. (2019). The influence of political knowledge on opinion polarization in citizen deliberation. *Political Research Exchange*, 1(1), 1–23. <https://doi.org/10.1080/2474736X.2019.1702887>
- Homana, G. A. (2018). Youth political engagement in Australia and the United States: Student councils and volunteer organizations as communities of practice. *Journal of Social Science Education*, 17(1), 41–54. <https://doi.org/10.4119/jsse-865>
- Hoskins, B., Saisana, M., & Villalba, C. M. H. (2015). Civic competence of youth in Europe: Measuring cross national variation through the creation of a composite indicator. *Social Indicators Research*, 123(2), 431–457. <https://doi.org/10.1007/s11205-014-0746-z>
- Ing M. (2018). What About the "Instruction" in Instructional Sensitivity? Raising a Validity Issue in Research on Instructional Sensitivity. *Educational and Psychological Measurement* 78(4), 635–652. <https://doi.org/10.1177/0013164417714846>
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, 34(1), 111–117. <https://doi.org/10.1177/001316447403400115>
- Kavadias, D., Hemmerechts, K., & Spruyt, B. (2017). Segregation and socialization: Academic segregation and citizenship attitudes of adolescents in comparative perspective? *Journal of Social Science Education*, 16(2), 30–41. <https://doi.org/10.4119/jsse-832>
- Keegan, P. (2021). Critical affective civic literacy: A framework for attending to political emotion in the social studies classroom. *The Journal of Social Studies Research*, 45(1), 15–24. <https://doi.org/10.1016/j.jssr.2020.06.003>
- Kenski, K., & Stroud, N. J. (2006). Connections between internet use and political efficacy, knowledge, and participation. *Journal of Broadcasting & Electronic Media*, 50(2), 173–192. https://doi.org/10.1207/s15506878jobem5002_1
- Kim, H. & Byun, S.-Y. (2019) Immigrant integration policy and native adolescents' attitudes towards ethnic minorities: a comparative study of European Countries, *Multicultural Education Review*, 11(3), 172–188. DOI: 10.1080/2005615X.2019.1644041
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., & Tenorth, H.-E. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. BMBF.
- Köller, O. (2016). Bildungsstandards. In R. Tippelt & B. Schmidt-Hertha (Eds.), *Handbuch Bildungsforschung* (pp. 625–648). Springer. https://doi.org/10.1007/978-3-531-20002-6_26-1
- Lange, H. (2002). PISA: Und was nun? *Zeitschrift für Erziehungswissenschaft*, 5(3), 455–471. <https://doi.org/10.1007/s11618-002-0061-3>

- Lauermann, F., Tsai, Y.-M., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy–value theory of achievement-related behaviors. *Developmental Psychology*, *53*(8), 1540.
- Lee, H. & Smith, W. Z. (2020). Fit Indices for Measurement Invariance Tests in the Thurstonian IRT Model. *Applied Psychological Measurement* *44*(4), 282-295. <https://doi.org/10.1177/0146621619893785>
- Levy, B. L. M., & Akiva, T. (2019). Motivating political participation among youth: An analysis of factors related to adolescents' political engagement. *Political Psychology*, *40*(5), 1039–1055.
- Liem, G. A. D., & Chua, B. L. (2013). An expectancy-value perspective of civic education motivation, learning and desirable outcomes. *Educational Psychology*, *33*(3), 283–313.
- Lührmann, A., & Lindberg, S. I. (2019). A third wave of autocratization is here: what is new about it? *Democratization*, *26*(7), 1095–1113. <https://doi.org/10.1080/13510347.2019.1582029>
- Massing, P. (2022). Kompetenzmodelle in der Politikdidaktik. In G. Weißeno & B. Ziegler (Eds.), *Handbuch Geschichts- und Politikdidaktik* (pp. 17–32). Springer. https://doi.org/10.1007/978-3-658-29668-1_2
- Maurissen, L. (2020). Political efficacy and interest as mediators of expected political participation among Belgian adolescents. *Applied Developmental Science*, *24*(4), 339–353. <https://doi.org/10.1080/10888691.2018.1507744>
- May, M. (2007). *Demokratiefähigkeit und Bürgerkompetenzen: Kompetenztheoretische und normative Grundlagen der politischen Bildung*. Springer. <https://doi.org/10.1007/978-3-531-90492-4>
- May, M., Moritz, B., & Riegel, P. (2020). Politische Urteilskompetenz empirisch – Überlegungen zur Operationalisierung und ein Test zur politischen Werturteilskompetenz. *Zeitschrift Für Didaktik Der Gesellschaftswissenschaften*, *11*(1), 134–146. <https://elibrary.utb.de/doi/10.46499/1601.1145>
- McDonnell, J. (2020). Municipality size, political efficacy and political participation: a systematic review. *Local Government Studies*, *46*(3), 331–350. <https://doi.org/10.1080/03003930.2019.1600510>
- McWhirter, E. H., & McWhirter, B. T. (2016). Critical consciousness and vocational development among Latina/o high school youth: initial development and testing of a measure. *Journal of Career Assessment*, *24*(3), 543–558. <https://doi.org/10.1177/1069072715599535>

- Messick, S. (1984). The Psychology of Educational Measurement. *Journal of Educational Measurement*, 21(3), 215–237.
- Meyer, J., Fleckenstein, J., & Köller, O. (2019). Expectancy value interactions and academic achievement: Differential relationships with achievement measures. *Contemporary Educational Psychology*, 58, 58–74. <https://doi.org/10.1016/j.cedpsych.2019.01.006>
- Ministry for Schools and Further Education of the State of North Rhine-Westphalia. (2022). *Das Schulwesen in NRW aus quantitativer Sicht 2021/22: Statistische Übersicht Nr. 417 [The school system in North Rhine-Westphalia from a quantitative perspective. 2016/17. Statistical overview nr. 417]*. MKW. https://www.schulministerium.nrw/system/files/media/document/file/quantita_2021.pdf
- Mondak, J. J. (2001). Developing valid knowledge scales. *American Journal of Political Science*, 224–238.
- Myoung, E. & Liou, P.-Y. (2022): Systematic review of empirical studies on international large-scale assessments of civic and citizenship education. *Scandinavian Journal of Educational Research*, 66(7), 1269–1291. <https://doi.org/10.1080/00313831.2022.2131903>
- National Assessment Governing Board. (2018). *Civics Framework: for the 2018 National Assessment of Educational Progress*. Washington, DC. National Assessment Governing Board, U.S. Department of Education. <https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/civics/2018-civics-framework.pdf>
- Oberle, M. (2022). Demokratiebildung in der Schule. In K. Möller, F. Neuscheler & F. Steinbrenner (Eds.), *Demokratie gestalten! Herausforderungen und Ansätze für Bildungs- und Sozialarbeit*. (pp. 62–73). Kohlhammer Verlag.
- Oberle, M. (2018). Politisches Effektivitätsgefühl von Schüler/-innen. Struktur, Determinanten und Veränderbarkeit einer motivationalen Facette politischer Kompetenz. In S. Manzel & M. Oberle (Eds.), *Kompetenzorientierung - Potenziale zur Professionalisierung der Politischen Bildung* (pp. 85–97). Springer.
- Oberle, M., & Wenzel, N. (2019). Politisches Vertrauen und Effektivitätsgefühl von Schüler/-innen - Einflussfaktoren und Relevanz für politische Partizipation. In M. Lotz & K. Pohl (Eds.), *Schriftenreihe der GPJE. Gesellschaft im Wandel: Neue Aufgaben für die politische Bildung und ihre Didaktik* (pp. 176–186). Wochenschau Verlag.
- OECD [The Organisation for Economic Cooperation and Development] (2001). *Knowledge and skills for life. First results from the OECD Programme for International Student Assessment (PISA) 2000*. OECD Publications.

- Otto, L., & Bacherle, P. (2011). Politisches Interesse Kurzsкала (PIKS): Entwicklung und Validierung. *Politische Psychologie*, *1*(1), 19–35.
- Prior, M. (2019). *Hooked: How politics captures people's interest*. Cambridge University Press.
- Prior, M., & Bougher, L. D. (2018). “Like they’ve never, ever seen in this country”? Political interest and voter engagement in 2016. *Public Opinion Quarterly*, *82*(S1), 822–842.
- Quintelier, E., & Hooghe, M. (2013). The relationship between political participation intentions of adolescents and a participatory democratic climate at school in 35 countries. *Oxford Review of Education*, *39*(5), 567–589. <https://doi.org/10.1080/03054985.2013.830097>
- Rasmussen, S. H. R., & Nørgaard, A. S. (2018). When and why does education matter? Motivation and resource effects in political efficacy. *European Journal of Political Research*, *57*(1), 24–46.
- Robitzsch, A., Kiefer, Thomas, Wu, & Margaret. (2021). *TAM: Test analysis module*. <https://cran.r-project.org/web/packages/TAM/index.html>
- Roccatò, M., & Zogmaister, C. (2010). Predicting the vote through implicit and explicit attitudes: A field research. *Political Psychology*, *31*(2), 249–274. <http://www.jstor.org/stable/20721288>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *2012*, *48*(2), 36. <https://doi.org/10.18637/jss.v048.i02>
- Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G., & Friedman, T. (2018). *Becoming citizens in a changing world: IEA International Civic and Citizenship Education Study 2016: International report*. Springer. <https://doi.org/10.1007/978-3-319-73963-2>
- Schulz, W., Fraillon, J., Losito, B., Agrusti, G., Ainley, J., Damiani, V., & Friedman, T. (2022). *International Civic and Citizenship Education Study 2022: Assessment Framework*. Springer.
- Stanley, B., & Czeńnik, M. (2022). Uninformed or informed populists? The relationship between political knowledge, socio-economic status and populist attitudes in Poland. *East European Politics*, *38*(1), 43–60. <https://doi.org/10.1080/21599165.2021.1876676>
- Stattin, H., & Amnå, E. (2022). Basic values transform political interest into diverse political values, attitudes and behaviors. *Journal of Youth and Adolescence*. Advance online publication. <https://doi.org/10.1007/s10964-022-01654-w>
- Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The importance of students’ motivation for their academic achievement — Replicating and extending

- previous findings. *Frontiers in Psychology*, 10, 1730. <https://doi.org/10.3389/fpsyg.2019.01730>
- van Deth, J. W. (1990). Interest in politics. In M. K. Jennings & J. W. van Deth (Eds.), *Continuities in political action* (pp. 275–312). De Gruyter. <https://doi.org/10.1515/9783110882193.275>
- Waldow, F. (2009). What PISA did and did not do: Germany after the ‘PISA-shock’. *European Educational Research Journal*, 8(3), 476–483. <https://doi.org/10.2304/eeerj.2009.8.3.476>
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Hogrefe & Huber Publishers.
- Weißeno, G., Detjen, J., Juchler, I., Massing, P., & Richter, D. (2010). *Konzepte der Politik - ein Kompetenzmodell. Schriftenreihe: Vol. 1016*. Bundeszentrale für politische Bildung. https://www.pedocs.de/volltexte/2016/12009/pdf/Weisseno_et_al_2010_Konzepte_der_Politik_.pdf
- Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, 34(5), 719–751. <https://doi.org/10.1177/0011000006286345>.
- Widmaier, B. & Nonnenmacher, F. (2011). Von der Politikverdrossenheit zum Wutbürger? Partizipation als Ziel der politischen Bildung – Zur Einführung. In B. Widmaier & F. Nonnenmacher (Eds.): *Partizipation als Bildungsziel. Politische Aktion in der politischen Bildung* (pp. 7–15). Wochenschau Verlag.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wise, S. L. (2017). Rapid-Guessing behavior: Its identification, interpretation, and implications. *Educational Measurement: Issues and Practice*, 36(4), 52–61. <https://doi.org/10.1111/emip.12165>
- Witschge, J. H. M. (2022). Education systems and inequality of civic and political engagement. [Doctoral thesis, Universiteit van Amsterdam]
- Wray-Lake, L., Metzger, A., & Syvertsen, A. K. (2017). Testing multidimensional models of youth civic engagement: Model comparisons, measurement invariance, and age differences. *Applied Developmental Science*, 21(4), 266–284. <https://doi.org/10.1080/10888691.2016.1205495>

Appendix A
Convergent and Discriminant Evidence, and Instructional Sensitivity
Table A1
Means of framework constructs in grade 7 and 10

Framework construct	<i>M</i> Grade 7	<i>M</i> Grade 10	Δ (<i>p</i>)
Political knowledge	0.03	0.75	0.72 (< .001)
Political interest	2.14	2.23	0.10 (.225)
Political efficacy	2.39	2.40	0.01 (.943)
Attitude towards democracy	2.92	3.14	0.22 (.020)
Attitude towards opinion pluralism	3.30	3.51	0.21 (.001)
Willingness to participate	2.42	2.48	0.06 (.467)

Note. *M* represents mean.

Table A2*Correlations within and across framework dimensions and with sense of social affiliation*

Framework construct	Within dimension		PI/ATD ^a		SOSA
Political knowledge	– ^b		.36	>	-.01 ^c
Political interest	.81	>	.38^c	>	.14^c
Political efficacy	.81	>	.50^c	>	.21^c
Attitude towards democracy	.39	>	.38	>	.14^c
Attitude towards opinion pluralism	.39	>	.18^c	<	.21
Willingness to participate	– ^b	>	.37^c	>	.17

Note. PI = Political interest; ATD = Attitude towards democracy; SOSA = Sense of social affiliation.

^aFor the cross-dimensional correlations, political interest was used. The exceptions were political efficacy and political interest itself, for which attitude towards democracy was used instead.

^bThe political knowledge dimension consisted of only one construct. ^cThe correlation differs from the correlation to the left of it at the 5% level of significance.

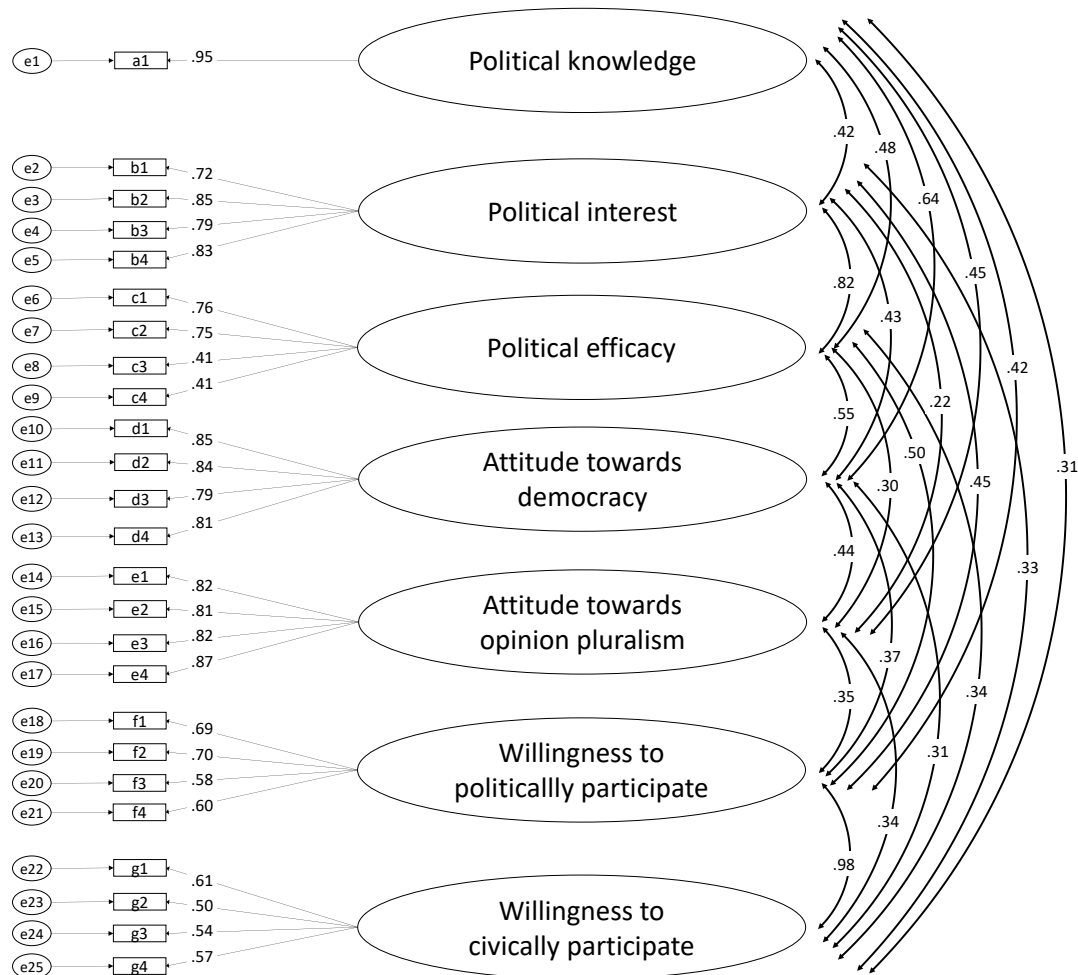
Correlations printed in bold reached statistical significance at the 5 % level.

Appendix B

Additional Confirmatory Factor Analyses

Figure B1

Confirmatory factor analysis (model 1)

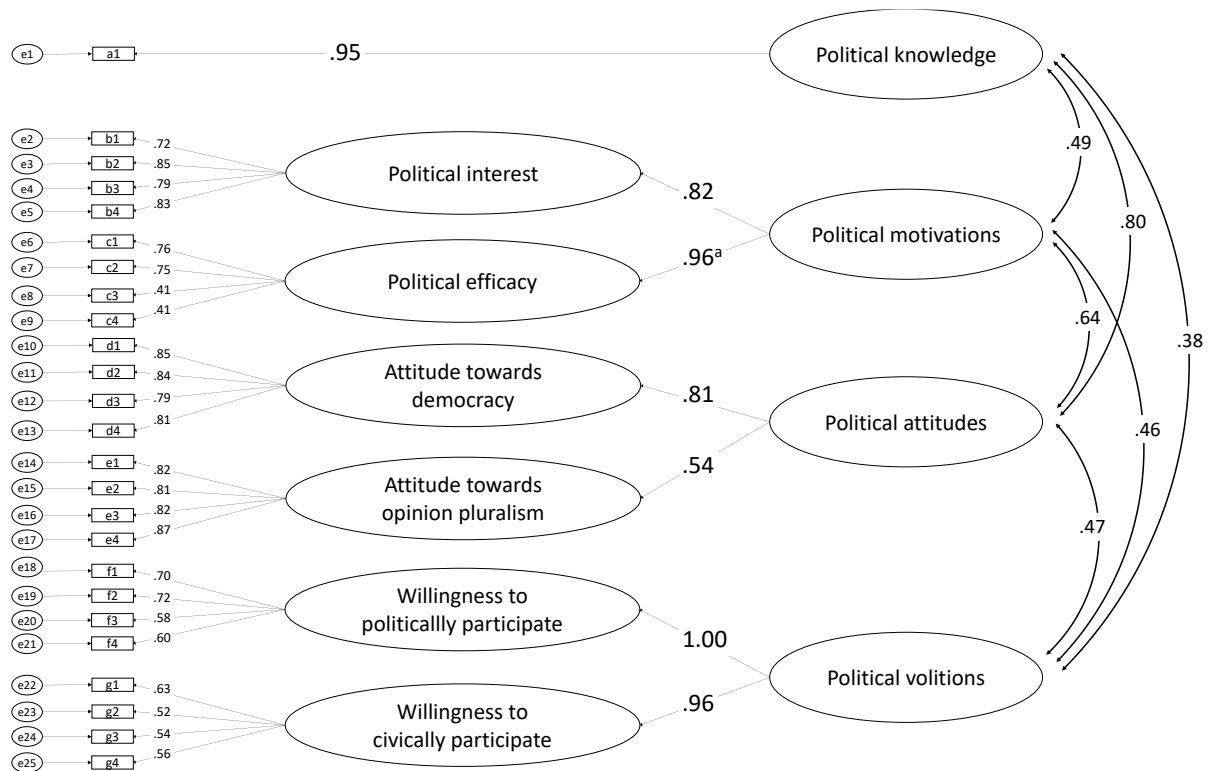


Note. CFI: .926; TLI: .913; RMSEA: .052; SRMR: .077.

All factor loadings and correlations reached statistical significance at the 5% level.

Figure B2

Confirmatory factor analysis (model 6)



Note. CFI: .923; TLI: .913; RMSEA: .052; SRMR: .080. When not reported otherwise, factor loadings/correlations reached statistical significance at the 5% level.

^aFactor loading/covariance did not reach significance.

Table B1*Comparison of Fit Indices in different CFAs*

Model	χ^2			Robust RMSEA		AIC	BIC	Δ AIC	Δ BIC
	Value	df	p^a	Value	90% CI				
M1: Facets only	989	256	–	.051	[.047, .055]	49121	49587	–	–
M6: Dimensions & facets	1029	265	< .000	.051	[.048, .055]	49144	49565	23	-22
Adjusted Model	1008	262	.004	.051	[.047, .055]	49129	49564	8	-23

Note. RMSEA = root-mean-square error of approximation; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.

^aThe χ^2 of Model 1 is used as the benchmark.

2.3 What Makes Domain Knowledge Difficult? Word Usage Frequency from SUBTLEX and dlexDB Explains Knowledge Item Difficulty

Ludewig, U., **Alscher, P.**, Chen, X. & McElvany, N. (2023). What makes domain knowledge difficult? Word usage frequency from SUBTLEX and dlexDB explains knowledge item difficulty. *Behavior Research Methods*, 55, 2621–2637. <https://doi.org/10.3758/s13428-022-01918-0>

Dieser Artikel stellt keine Kopie der Veröffentlichung dar und entspricht möglicherweise nicht exakt der endgültigen und maßgeblichen Version des in der Zeitschrift veröffentlichten Artikels.

Abstract

The quality of tests in psychological and assessment educational is of great scholarly and public interest. Item difficulty models are vital to generating test result interpretations based on evidence. A major determining factor of item difficulty in knowledge tests is the opportunity to learn about the facts and concepts in question. Knowledge is mainly conveyed through language. Exposure to language associated with facts and concepts might be an indicator of the opportunity to learn. Thus, we hypothesize that item difficulty in knowledge tests should be related to the probability of exposure to the item content in everyday life and/or academic settings and therefore also to word frequency. Results from a study with 99 political knowledge test items administered to $N = 250$ German 7th (age: 11 – 14 years) and 10 th (age: 15 – 18 years) graders showed that word frequencies in everyday settings (SUBTLEX-DE) explain variance in item difficulty, while word frequencies in academic settings (dlexDB) alone do not. However, both types of word frequency combined explain a considerable amount of the variance in item difficulty. Items with words that are more frequent in both settings and, in particular, relatively frequent in everyday settings are easier. High word frequencies and relatively higher word frequency in everyday settings could be associated with higher probability of exposure, conceptual complexity and better readability of item content. Examining word frequency from different language settings can help researchers investigate test score interpretations and is a useful tool for predicting item difficulty and refining knowledge test items.

Keywords: Political knowledge · Word frequency · Item difficulty · Statistical suppression effect · Educational assessment

The design of high-quality assessments for knowledge, abilities, and competencies is a major research topic in educational assessment and psychometrics (APA, 2020; Care et al., 2018). A high-quality assessment should be based on a solid theory about the domain, and this theory should be able to explain why items are difficult or easy (Mislevy et al., 2003). Essentially, difficulty is a property of an item that describes how much skill, ability, or knowledge is required to solve the item (Embretson & Reise, 2013). Domain-related and theory-based features of test items should explain item difficulty to allow valid interpretations of test results. Different approaches to identifying domain-related item features have proven successful in various fields. One way to organise these approaches is to divide them into structure-driven, complexity-driven and exposure-driven approaches.

Structure-driven approaches assume that a domain consists of distinct elements (i.e., concepts or skills) that have a defined relationship to one another, and solving an item requires some subset of these elements. This approach has been formally defined in (probabilistic) knowledge space theory (e.g., Stefanutti et al., 2012). For example, Tatsuoka (1990) described the domain of solving fractions problems on the basis of seven elements, termed skills (e.g., distinguishing whole numbers from fractions or converting whole numbers to fractions). The domain structure of fractions is hierarchical, because some skills are prerequisites for other skills (e.g., performing the basic fraction subtraction operation and distinguishing whole numbers from fractions are prerequisites for borrowing one from the whole number to the fraction). Other examples of concise knowledge structures can be found in stoichiometry (mathematical chemistry; Segedinac et al., 2018), stochastic problem solving (Stefanutti et al., 2012) or the laws of mechanics (Reif & Heller, 1982). In concise hierarchical domain structures, items are characterized based on the specific subset of skills required, students fail to answer items if they lack a skill, and items that require more and higher-order skills are more difficult.

Complexity-driven approaches assume that proficiency relates to processing capacity and item difficulty to complexity. Broadly defined, complexity refers to a number of variable elements that must be related to each other to answer an item. In spatial (e.g., Embretson & Yang, 2006) and analogical reasoning (Stevenson et al., 2013), the cognitive complexity of items is defined based on the number and type of cognitive operations (e.g., mentally rotating or mirroring shapes) necessary to relate all variable elements and to falsify or verify response options. In passage comprehension, one measure of complexity is propositional idea density, which assesses the number of propositions that need to be related to answer an item (e.g., Ozuru et al.,

2008). In complexity-driven approaches, items are characterized via additive features that contribute to complexity, students fail to answer items because the complexity exceeds their processing capacity, and items that are more complex are more difficult.

Exposure-driven approaches characterize items based on the frequency and intensity of associated learning opportunities. For word recognition, items with infrequent words are more difficult to solve than items with frequent words (e.g., Brysbaert et al., 2019). Politicians who are present in the media are easier to know than ones who are less present (Westle & Tausendpfund, 2019). Overall presence of information has been found to be a main driver for item difficulty in knowledge tests about authors, newspapers and television (see the environmental opportunity hypothesis; Stanovich & Cunningham, 1993). Thus, items can be characterized according to measures of exposure frequency or presence; students fail to answer items because they were insufficiently exposed to the underlying content or lack the ability to learn from exposure. Items with content for which there are few opportunities to learn should be more difficult.

In educational psychology and assessment, domain knowledge tests are important for theory development (e.g., Kim et al., 2021) and monitoring educational outcomes (National Research Council, 2012). However, relatively few studies have systematically examined item difficulty on domain knowledge tests. We define domain knowledge as factual (e.g., knowledge of terminology) and conceptual knowledge (e.g., knowledge of theories, models, and structures) relevant to a particular domain (e.g., science or policy). In general, it is plausible to assume that structural features (e.g., some concepts might be on a higher order than others), complexity features (e.g., some concepts might be inherently more complex than others), and exposure features (e.g., some concepts might be more present than others) influence item difficulty in domain knowledge tests.

There are models that describe the structure and complexity of domain-specific knowledge (e.g., science: Kim et al., 2021; politics: Weißeno et al., 2010) and more holistic, domain-general taxonomies (e.g., types and qualities of knowledge: De Jong & Ferguson-Hessler, 1996; Bloom's taxonomies: Krathwohl & Anderson, 2010). However, these models usually do not state what particular factual or conceptual entities are difficult or easy (e.g., Tauber et al., 2013). Thus, it seems worthwhile to investigate objective indicators explaining the difficulty of items in domain knowledge tests.

The current study developed a new approach to explain the difficulty of items in knowledge tests, focusing on factual and conceptual knowledge in the political domain. Knowledge is mainly conveyed through language; therefore, language use could be an indicator

of knowledge item difficulty. Using corpus databases from everyday and academic settings, it is possible to measure how frequently words are used in a given setting. Word frequency could be indicative of knowledge item difficulty because usage frequency could be related to the likelihood of being exposed to the item content, and relative frequency in everyday settings could be associated with real-life experiences. Word frequency has rarely been applied to assess educational achievement, likely due to a lack of understanding of the relationship between domain knowledge and language use. The present study shows that word frequency can be useful for predicting item difficulty in domain knowledge tests.

Political Knowledge and Language

Political knowledge involves being able to recall from memory facts about a political system (i.e., terminology, theories, or structures) that are relevant for interpreting and understanding happenings and developments within that system (Clark, 2017). As such, political knowledge helps people understand political debates and their relevance, sort and categorize political information, and become aware of their own political needs and preferences and what political actions and decisions must be pursued to satisfy these needs and preferences (Cramer & Toff, 2017).

Political knowledge is mainly conveyed and expressed through language. Domain knowledge can be acquired intentionally through academic activities in school or incidentally through exposure to media and real-life experience (Irwing et al., 2001). Students learn factual and conceptual knowledge about political issues through media consumption (i.e., news, television, radio and social media: Bischof & Senninger, 2018) as well as through oral and written discourse in everyday life or in the school context (e.g., Carpini & Keeter, 1996). There are different ways to learn political facts and concepts, but they all are primarily conveyed through language.

We hypothesize that two major aspects might influence the difficulty of political knowledge items and items in other knowledge domains tightly linked to language use. First, how present is the particular topic in people's lives? Most people know more about present and salient issues (i.e., that people speak, hear and read about) in their lives (e.g., what democratic institution makes laws that directly affect people's lives) than those that are rarer (e.g., certain laws that only apply in exceptional situations).

Second, in what settings are these issues present? Some are more present in everyday life (i.e., knowing about one's country's current head of government) than in academic settings.

Other issues are more present in academic settings (i.e., the structure of the separation of powers) than in everyday life.

In language research, word frequency has been considered an indicator for the probability of exposure and used to explain the difficulty of test material. It is also considered important to distinguish between different language use settings.

Word Frequency and Exposure

Word usage frequencies have impressive explanatory power for many language-related tasks. Visual word recognition is slower and more likely to be incorrect when words are used infrequently (Brysbaert et al., 2018). Word frequency is associated with accuracy and latency in semantic classification tasks (Taikh et al., 2015). Infrequent words are more often unknown than frequent words (Brysbaert et al., 2019). Texts are more challenging if they contain rarer words (Berendes et al., 2018; Fitzgerald et al., 2015) and word frequency calculated based on different corpora contribute to explaining text complexity (Chen & Meurers, 2018).

There is a complex debate on the underlying causes of the word frequency effect (see Brysbaert et al., 2018, for a more detailed review). One primary and quite intuitive reason for the word frequency effect is that individuals are frequently exposed to words in language use. With repeated exposures, words become more accessible, and it is more likely that individuals associate a distinct meaning with them (Juhasz et al., 2019). Furthermore, research shows that word frequency based on corpora is most representative of everyday language settings and explains performance in word recognition tasks better than word frequencies gathered from less representative settings (Brysbaert et al., 2011). Individuals are exposed to frequent words more often, and the probability of word exposure is associated with the probability and extent of being familiar with words.

Word exposure and exposure to concepts and facts in the political domain are not inevitably connected, but should be associated with one another in authentic situations. On the one hand, students could theoretically be frequently exposed to the word “parliament” in some de-contextualized way (e.g., in spelling excesses). Thus, word exposure need not necessarily be an opportunity to learn about the concept of a parliament. On the other hand, concepts and facts about parliaments could be conveyed using a synonym (e.g., “congress” or “legislature”). Thus, exposure to a specific word is not a necessary prerequisite for learning about the concept of a parliament. However, in authentic language use, exposure to words will be embedded in a related context, and learning about concepts will probably involve exposure to different synonym

words. In authentic language use contexts, a test item's word frequency could therefore be a good indicator of exposure to associated concepts and facts.

Language Setting

Exposure in everyday language settings might not be the only relevant aspect related to word frequency. Knowledge is passed on and expressed through academic language. Academic language is the specialized language in academic settings that facilitates communication and thinking about specific content domains (Nagy & Townsend, 2012). A good representation of language in academic settings are texts that convey knowledge, such as textbooks, lexicons, or encyclopaedias (Coxhead, 2000).

In contrast to word frequency in everyday language settings, word frequency could potentially influence knowledge test item difficulty in two ways. On the one hand, words that are frequent in these academic settings are helpful for communicating knowledge and should be more familiar to students who have been exposed to academic content. Thus, words frequent in academic settings could be an indicator for probability of exposure in academic contexts. However, students are much less exposed to language in academic than in everyday settings. Therefore, word frequency in academic settings might be an inferior indicator for the probability of exposure compared to word frequency in everyday language settings (Brysbaert et al., 2011; Coxhead, 2000).

On the other hand, words that are frequent in academic settings might be associated with more academically cultivated language, used in formal definitions of terminology, theories, models, and structures. Thus, controlling for word frequency in everyday settings (i.e., the best indicator for exposure), word frequency in academic settings could explain additional variance in item difficulty by capturing the degree to which the item content is divorced from everyday language use.

In sum, word frequencies could be beneficial for determining difficulty in knowledge tests, such as those for political knowledge because they could be an indicator for probability of exposure to the item content and/or the degree academic cultivation. Presumably, one of the best indicators of exposure is word frequency in everyday life settings. Word frequency in language settings used to convey knowledge (i.e., academic language) could additionally contribute to explaining item difficulty. On the one hand, it could be a congruent indicator of exposure in academic settings; on the other hand, it could be a complementary indicator that captures the content's degree of formal and academic sophistication as indicated by its divergence from word frequency in the everyday setting.

The Present Study

Understanding what features make domain knowledge tests difficult is vital for supporting an evidence-based test result interpretation. Word frequency can be viewed as a proxy for exposure probability. Word frequency might be an indicator for the frequency of opportunities to learn about political concepts and facts in everyday life and academic settings. The frequency of opportunities to learn should be a main driver of knowledge item difficulty.

Considered separately, does word frequency in everyday and academic settings have a congruent effect on explaining the difficulty of items in political knowledge tests?

- H1: Average word frequency in everyday and academic settings individually explain item difficulty
 - H1a: Average word frequency in everyday settings is negatively associated with item difficulty (Low word frequency is associated with more difficult items)
 - H1b: Average word frequency in academic settings is negatively associated with item difficulty (Low word frequency is associated with more difficult items)

Considered together, does combining word frequencies from everyday and academic settings in one analysis have a complementary effect?

- H2: Combining average word frequency from everyday and academic settings leads to a significant increase in the explained variance in item difficulty.

Methods

Participants

Seventh- and 10th-grade students from German schools in middle-sized cities participated in a study on the development of political and civic competencies among youth in fall 2019. In total, 152 7th graders ($M_{age} = 12.54$, $SD = 0.91$, range = 11–14 years; 45% female) and 98 10th graders ($M_{age} = 16.12$, $SD = 0.97$, range = 15–18 years; 35% female) participated in the study. The sample included students from four types of German public schools, ‘Hauptschulen’ (lower vocational track), ‘Realschulen’ (vocational track), ‘Gymnasium’ (academic track) and ‘Gesamtschulen’ (comprehensive schools). Among participants, 43.6% had an immigrant background (i.e., either one or both parents were born outside of Germany). The proportion of students from immigrant backgrounds is higher than the national average in Germany but common for urban West Germany. Additionally, the proportion of females in the tenth grade was significantly below 50% because one of the 10th-grade schools had a vocational orientation that

primarily attracted male students. Thus, the sample captures the variability in achievement levels among German students but is not representative. However, unbiased estimates of item difficulties can be derived from unrepresentative samples (Embretson & Reise, 2013; implications discussed in the limitations). The responsible Ethics Committee approved the study, and all participants or their parents (if the students were younger than 16) were asked to give their informed consent. Only participants with valid informed consent forms were allowed to participate in the study.

Materials

Political Knowledge Test

Items. The political knowledge assessment was developed for a national large-scale assessment study and included 99 items covering different aspects and facets of political knowledge (Alscher et al., 2022). The items were constructed in a workshop with five content experts and one test administration expert. The items were initially constructed by a team led by a political scientist; then, the experts reviewed all items independently in terms of factorial correctness, solvability, and grade appropriateness. Additionally, the experts revised the items independently regarding language use. The goal was to create items with suitable and authentic language that is as simple as possible and as complex as necessary.

Figure 1

Example items in original German (top) and English (bottom)

<p>Der Begriff "Europäische Integration" bezeichnet die Zusammenarbeit und Erweiterung der Europäischen Union. Welche der folgenden Aussagen ist <u>kein</u> Argument für die Zusammenarbeit innerhalb Europas?</p> <p>Bitte wähle <u>eine</u> der folgenden Antworten.</p> <ul style="list-style-type: none"> <input type="radio"/> Durch die politische Zusammenarbeit kann sich Europa gegenüber anderen Ländern besser behaupten. <input type="radio"/> Durch die militärische Zusammenarbeit wird die Gefahr von Konflikten und Kriegen innerhalb Europas geringer. <input type="radio"/> Durch die wirtschaftliche Zusammenarbeit wird der Handel zwischen europäischen Ländern vereinfacht. <input type="radio"/> Durch die gesellschaftliche Zusammenarbeit werden nationale Regeln und Gesetze weniger wichtig. 	<p>Bald ist wieder Bundestagswahl. Angenommen du bist alt genug, um wählen zu gehen. Worauf sollte deine Wahlentscheidung, im Sinne einer Demokratie, <u>bestenfalls</u> basieren?</p> <p>Ich wähle die Partei, die ...</p> <p>Bitte wähle <u>eine</u> der folgenden Antworten.</p> <ul style="list-style-type: none"> <input type="radio"/> meine Eltern wählen. <input type="radio"/> wahrscheinlich die Wahl gewinnt. <input type="radio"/> am besten meine Standpunkte vertritt. <input type="radio"/> ich am häufigsten gesehen habe.
<p>The term "European integration" refers to the cooperation and expansion of the European Union. Which of the following statements is <u>not</u> an argument for cooperation within Europe?</p> <p>Please select <u>one</u> option</p> <ul style="list-style-type: none"> <input type="radio"/> Through political cooperation, Europe can better assert itself against other countries. <input type="radio"/> Military cooperation reduces the risk of conflict and war within Europe. <input type="radio"/> Economic cooperation makes trade between European countries easier. <input type="radio"/> Social cooperation makes national rules and laws less important. 	<p>Soon there will be another federal election. Suppose you are old enough to vote. From a democratic perspective, what should your voting decision <u>ideally</u> be based on? I vote for the party that ...</p> <p>Please select <u>one</u> option</p> <ul style="list-style-type: none"> <input type="radio"/> my parents vote for. <input type="radio"/> is probably going to win the election. <input type="radio"/> best represents my points of view. <input type="radio"/> I have seen most often.

Note. Item 51 (left) and 29 (right).

The 99 items included 36 grade-specific items and 27 anchor items that students from both grades answered. All items are multiple-choice with four answer options, only one of

which is correct. The items consist only of text (i.e., no figures or tables) and consist of between 22 and 154 words (prompt, question, answer options). Overall, the test had a high reliability $REL_{cap} = .89$. All items combined included a total of 4,819 words (token) and 1,569 types (individual or unique words).

Item Difficulty. Item difficulty is an item trait given by the proportion of students who are capable of answering the item correctly. We applied an effort moderated (Wise & DeMars, 2006) unidimensional Rasch multi-group IRT model with the TAM package (Robitzsch et al., 2018) in R (R Core Team, 2014) using marginal maximum likelihood to estimate the item difficulty parameter. Instead of point estimates, we used a plausible value approach with ten drawings to enable a measurement error-adjusted and unbiased estimation of effects in the further analysis.

Word Frequency

Word frequency in everyday settings was measured with SUBTLEX-DE (<http://crr.ugent.be/archives/534>; Brysbaert et al., 2011). SUBTLEX-DE consists of German-language subtitles from 4,610 films and television shows, resulting in a corpus of 25,399,040 words (i.e., tokens) and 319,536 different words (i.e., types). In total, only 8.41% of all types in the items were not found in SUBTLEX-DE.

Word frequency in academic settings was measured via the lexical database dlexDB. DlexDB is based on the core corpus of the Digital Dictionary of the German Language (DWDS). The DWDS core corpus is a reference corpus of the German language in the 20th century, balanced in terms of time and text types, and has the following composition in its online version (fiction: ca. 28%, newspapers: ca. 27%, academic literature: ca. 23%, practical texts: ca. 21%). The core corpus of the DWDS has a volume of ca. 100 million running text words (tokens). The number of different words (types) is approximately 2.3 million. In total, only 1.91% of all types in the items were not found in dlexDB.

We used multiple imputation to address the missing word frequencies. Multiple imputation has been shown to be the least biased method of dealing with missing data (Sinharay et al., 2001). Multiple imputation is most effective when missing values can be imputed based on non-missing information; therefore, it is best practice to include additional variables, that are not part of the intended analysis (also called “auxiliary variables”: Mustillo, & Kwon, 2015). For the multiple imputation, (1) we gathered auxiliary variables: word frequencies from, web texts (541,453,764 tokens; 6,303,178 types; non-found: 0.32%), German Wikipedia in 2021 (<https://wortschatz.uni-leipzig.de/de>; 17,765,613 tokens; 983,883 types, non-found: 2.42%),

the Part-of-Speech, and character length, (2) we used the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011), the *norm* method (i.e., imputation via Bayesian linear regression), 200 restarts, and 20 imputed datasets. Multiple imputation should increase the reproducibility of our results (e.g., in other languages) because the variance of items' average word frequencies and the covariance between word frequencies from different corpora will be less biased by corpus size than with alternative methods.

We computed so-called Zipf values based on the raw word frequencies with capitalization normalization (Diependaele et al., 2013; Van Heuven et al., 2014). Zipf values are logarithmically scaled, account for the size of a corpus, and are transformed so that a value of 3 corresponds to the frequency of a word that occurs once in a million words, a value of 4 ten times in a million words, a value of 5 100 times in a million words, etc. Finally, we calculated the arithmetic average of all word (i.e., type) frequencies for the analysis on the item level.

We calculated three different average word frequencies. First, a simple average word frequency including all words in items. Second, the mean frequency of all words except stop words, using the *R* package *stopwords* (Benoit et al., 2021), to decrease the effect of very frequent function words. Third, we calculated the mean frequency of nouns, verbs and adjectives because nouns, verbs and adjectives should best characterize the actual facts and concepts addressed in an item. For the presented analysis, we primarily report the average frequency of nouns, verbs and adjectives (additional analyses can be found in Appendix B).

Procedure

The political knowledge assessment was administered as part of a study on civic literacy on 10.1" tablets in class settings. The political knowledge assessment was the first part of the study and took 60 minutes. Each student completed the 27 anchor items and 36 grade-specific items. The items were presented in nine different orders. The item orders were permuted block-wise in a Latin square to counterbalance order effects (Frey et al., 2009). The blocks included equal proportions of items from different content areas and anchor items. The test was administered with a forced-choice answer format. To reduce the influence of rapid-guessing behavior, we visually identified the threshold in the response time distribution (following the recommendations of Wise & DeMars, 2006) at 4.5 seconds, leading to the deletion of 2.48% of responses. In the end, together with not reached items, a total of 5.71% of all responses were missing. The full study took 3 hours. Besides the knowledge assessment, students were asked to answer demographic and political orientation questions.

Analysis

Presented Analysis

For the analysis, first, we used an ordinary least squared regression (OLR) using the *lm* function from the *base* package (R Core Team, 2014) to explain the item difficulty parameters based on the (1) average word frequency in everyday settings alone, (2) average word frequency in academic settings alone, and (3) both word frequencies simultaneously. All results are based on coefficients pooled from 20 consecutive analyses using the 20 imputed datasets with plausible item difficulty values. Second, we calculated the significance of changes in (pooled) coefficient size between 1 and 3 or 2 and 3, which indicate mediation or suppression effects, respectively, according to MacKinnon et al. (2000). Third, we present a regression based on orthogonal principle component analysis (PCR) using the *prcomp* function from the *stats* package (R Core Team, 2014).

According to G*Power (Faul et al., 2007), given 99 observations (items), p -value $p < .05$, two predictors and 80% test power, the analysis has an ability to detect a medium effect $f^2 \geq .10$ ($R^2 = .091$). For the R^2 increase ($p < .05$, Power = 80%) from a one-predictor model to a two-predictor model, the detectable effect is $f^2 \geq 0.081$ ($partialR^2 = .074$). The analysis was not preregistered. The data and scripts are available at <https://osf.io/bsn9m/>.

Robustness Analysis

Word frequencies in different settings are highly correlated, and under some conditions, multicollinearity can cause computational problems (Cohen et al. 2003). First, we calculated the variance inflation factor (VIF) for the combined model. The VIF ranged between $VIF = 2.91$ and 3.10 and thus did not indicate problematic multicollinearity (critical $VIF > 5$; Akinwande et al., 2015). Second, we validated that the correlation matrix is non-negatively defined (Friedman & Wall, 2005). Third, we replicated the results using resampling methods with the *train* function from the *caret* package (Kuhn, 2021), applying the method “repeatedcv” (with ten repetitions of 5-fold cross-validation) and “boot632” (with 1000 bootstraps). The regression weights and R^2 estimates from the resampling methods did not deviate from the original OLR. Fourth, we applied a ridge regression using the *glmnet* function from the package with the same name (Friedman et al., 2010) to validate the R^2 of the combined model. The ridge regression did not yield a different R^2 . Fifth, we replicated the estimates with the formula presented by Friedman & Wall, 2005 (formula and results can be found in Appendix C). The results did not deviate from the OLR. Sixth, we calculated the results using residualized variables for academic and everyday word frequency (Wurm & Fisicaro, 2014; formula and results can be found in

Appendix C). The robustness analyses did not yield different estimates or different interpretations than the presented analysis.

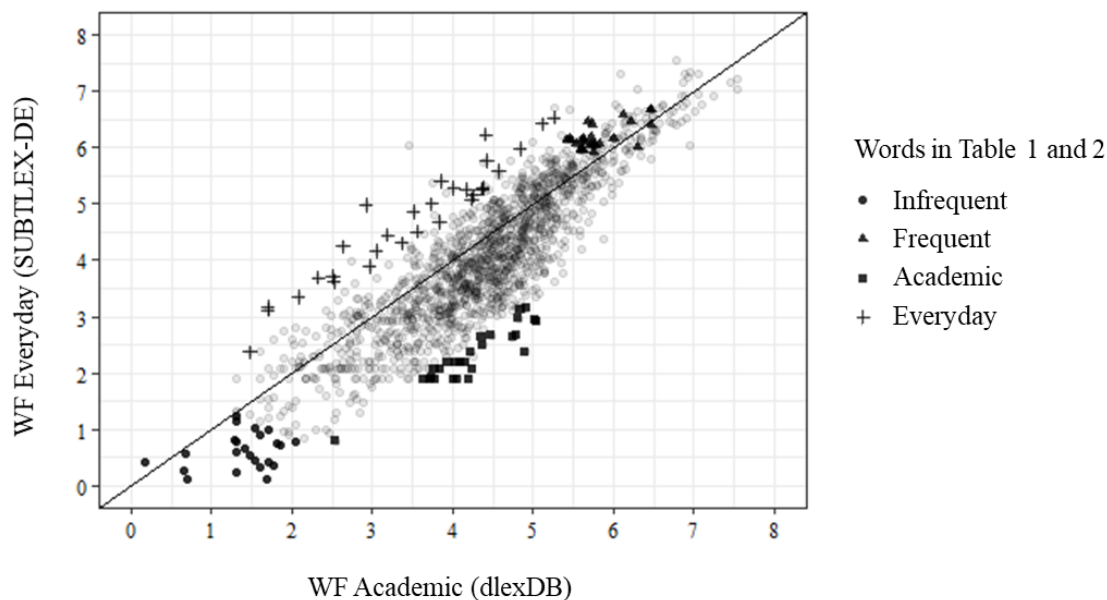
Results

Preliminary Results

Figure 2 shows the bivariate distribution of word frequency in academic and everyday settings. The two word frequencies are highly correlated, $r(1, 567) = .82$ (without stop words: $r = 0.76$; only nouns, verbs and adjectives: $r = .77$).

Figure 2

Bivariate distribution of word frequency in everyday and academic settings



Note. Words above the diagonal are relatively more frequent in academic settings and words below the diagonal are relatively more frequent in everyday settings.

To illustrate what words are frequent, infrequent, and relatively more frequent in everyday settings, and relatively more frequent in academic settings, we display the top 30 words for each setting in Table 1 and Table 2. The most frequent words are mostly auxiliary verbs. Words that are relatively more frequent in everyday settings include colloquial terms (e.g., *bescheuert* [stupid] or *entschuldigen* [saying sorry]) and terms that are related to students' life situation (e.g., *Schulabschluss* [graduation], *Klassenstufen* [grade] or *Vorstellungsgespräch* [job interview]). The infrequent words are mostly nouns specific to political topics. Many of them are composite words. Words that are relatively more frequent in academic settings encompass

fewer composite words, more key terms related to the topic of politics (e.g., Bundesrepublik [Federal Republic]).

Table 1

Top 30 most frequent (see Figure 2; triangles) words (nouns, verbs, and adjectives) and most infrequent words (circles)

	Frequent		Infrequent	
	German	English	German	English
1	hat	has	Regierungsvorsitzenden	Government Chair
2	sind	are	Ausgangsbeschränkungen	Output restrictions
3	wird	will	Umweltvereins	Environmental association
4	habe	have	Arbeitnehmerschutzes	Worker protection
5	kann	can	Spendenquittungen	Donation receipts
6	hatte	had	Supranationalitätsprinzip	Supranationality principle
7	können	can	Pflichtversicherungen	Compulsory insurance
8	meine	my	Nationenprinzip	Principle of nations
9	gut	well	Hoheitsprinzip	Sovereignty principle
10	machen	make	Konfliktprinzip	Conflict principle
11	soll	should	Gleichbehandlungsgesetz	Equal treatment law
12	gibt	gives	Asylberechtigung	Right of asylum
13	leben	live	Fachministerin	Specialized minister
14	müssen	must	Kollegialitätsprinzip	Principle of collegiality
15	sagen	say	Gesetzesvorschlägen	Legislative proposals
16	viel	much	Kanzlerprinzip	Chancellor principle
17	geht	goes	Ressortprinzip	Departmental principle
18	hast	have	Menschenrechtsverstöße	Human rights violations
19	vielleicht	maybe	Klassensprechern	Class representatives
20	lassen	let	Eu-ebene	Eu level
21	kommt	comes	Anführerinnen	Leaders
22	tun	do	Bundeländern	Federal states
23	gehen	go	Diktatorin	Dictator
24	wissen	know	Verursacherinnen	Causers
25	macht	power	Verbrecherinnen	Criminals
26	kommen	come	Luxusreisen	Luxury travel
27	wollte	wanted	Politikunterricht	Politics lessons
28	wollen	want	Strafverfolgungen	Prosecutions
29	wirklich	really	Anwohnerinnen	Residents
30	bist	are	Schulsprecher	Head boy

Note. We classify the words into most frequent, most infrequent, academic and everyday lists to provide illustrative examples. The classification had no influence on the further analysis and is not mutually exclusive (e.g., words in the everyday list can be in the most frequent list).

Table 2

Top 30 words (nouns, verbs, and adjectives) more frequent in everyday settings (see Figure 2; crosses) and more frequent in academic settings (squares)

	Everyday		Academic	
	German	English	German	English
1	passt	fits (slang)	Bundesrepublik	Federal Republic
2	muss	must	Parlamentarischen	Parliamentary
3	bescheuert	stupid	Regierungsvorsitzenden	President of the Government
4	solltest	should	Staatspräsident	President of the Republic
5	Schulabschluss	School graduation	Bundeskanzlers	Chancellor (possessiv)
6	Klassensprecher	class president	Deutschlands	Germanys
7	Reporterin	reporter	Grundgesetzes	Constitution (possessiv)
8	müsste	should	Regelung	Regulation
9	passiert	happens	Bundeskanzler	Federal Chancellor
10	bist	are	wirtschaftlichen	economic
11	lässt	lets	Wirtschaftspolitik	economic policy
12	hast	have	bayerische	Bavarian
13	Klassenstufen	grades	Prinzipielle	Principle
14	möchtest	would like	bestehende	existing
15	müssten	would have to	Entfaltung	development
16	bittest	ask	Gesellschaftlichen	Social
17	kannst	can	Herrschende	Ruling
18	abschalten	switch off	Zielsetzungen	Objectives
19	Vorstellungsgespräch	job interview	Spendenquittungen	Donation receipts
20	entschuldigen	sorry	Bundesverfassungsgericht	Federal Constitutional Court
21	stimmt	true	Institutionen	Institutions
22	Anführer	leader	Pflichtversicherungen	Compulsory insurance
23	verschwenden	waste	Neuwahlen	New elections
24	Chefin	boss	Bundesregierung	Federal government
25	gefällt	like	hochgebildeter	highly educated
26	kriegen	get	Verflechtung	interconnectedness
27	Telefon	Phone	Sozialpolitik	Social policy
28	Polizistinnen	policewomen	Solidarität	solidarity
29	passieren	pass	wirtschaftliche	economic
30	aufpassen	watch	Völkerrecht	international law

Note. We classify the words into most frequent, most infrequent, academic and everyday lists to provide illustrative examples. The classification had no influence on the further analysis and is not mutually exclusive (e.g., words in the everyday list can be in the most frequent list).

Descriptive Results

Average word frequency was $M = 4.15$, $SD = 0.35$ in everyday settings and $M = 4.49$, $SD = 0.26$ in academic settings. Average word frequency in the two settings was highly correlated, $r(97) = .77$, $p > .001$. Across different measures, the average word frequency in everyday settings was significantly correlated with item difficulty, $r(97) = -.36$, $p < .001$, whereas word frequency in academic settings was not, $r(97) = -.09$, $p = .352$ (see Table 3).

In addition, we performed a principal component analysis to obtain two orthogonal components. The first principal component was positively correlated with everyday and academic word frequency, $r(97) = .94, p < .001$, so we called it *shared frequency*. The second principal component was positively correlated with everyday word frequency, $r(97) = .34, p = .001$, and negatively correlated with academic word frequency, $r(97) = -.34, p = .001$, so we termed it *everydayness*. Shared frequency, $r(97) = -.23, p = .023$, and everydayness, $r(97) = -.36, p < .001$, were negatively correlated with item difficulty.

There were slight differences in average word frequency including all words, without stop words and only considering nouns, verbs and adjectives with regard to the mean, standard deviation, and correlations that were not statistically significant. The overall average word frequency was lowest when stop words were excluded. The standard deviation was highest when we only considered nouns, verbs and adjectives (results can be found in Appendix B).

Table 3

Correlations and descriptive statistics

Variables	1	2	3	4	5
1. WF everyday					
2. WF academic	.77^a				
3. Shared frequency	.94^a	.94^a			
4. Everydayness	.34^a	-.34^a	0 ^a		
5. Item difficulty ^a	-.34^b	-.09^b	-.25^c	-.36^c	
<i>M</i>	4.15	4.49	0	0	0
<i>SD</i>	0.35	0.26	1.33	0.48	1.18

Note. $N = 99$ Political knowledge items. Bold correlations, $p < .05$. WF: Average word frequency of nouns, verbs and adjectives. Light grey fields: Correlations with item difficulty.

^aPlease find histogram displaying the distribution of item difficulty in Appendix A. Comparison of correlations from dependent samples: ^bCorrelations significantly different, $z = 9.19, p < .001$,

^cCorrelations not significantly different, $z = 1.51, p = .065$.

Does word frequency in academic and everyday settings explain item difficulty in political knowledge assessment?

We present the regression results for average word frequency using only nouns, verbs, and adjectives in the results section. Please find the equivalent analyses for average word frequency with all words and without stop words in Appendix B Table B1. The core findings do not deviate between different average word frequencies.

Considered separately, is there a congruent Effect?

Word frequency in everyday settings had a statistically significant effect on item difficulty, $\beta_1 = -0.34$, $t = -3.51$, $p < .001$. Items become harder as word frequency in everyday settings drops, supporting H1a. Word frequency in everyday settings explains 13% of the variance in item difficulty. In contrast, word frequency in academic settings has no significant effect, $\beta_2 = -0.09$, $t = -0.91$, $p = .364$. Thus, Hypothesis H1b is not supported by the analysis.

Considered together, is there a complementary effect?

We evaluate the relative contribution of average word frequency from everyday and academic settings to explaining difficulty by entering them both in the regression analysis. In the combined model, both word frequencies have an effect on item difficulty (see Table 4).

Table 4

Regression analysis explaining item difficulty in the political knowledge test

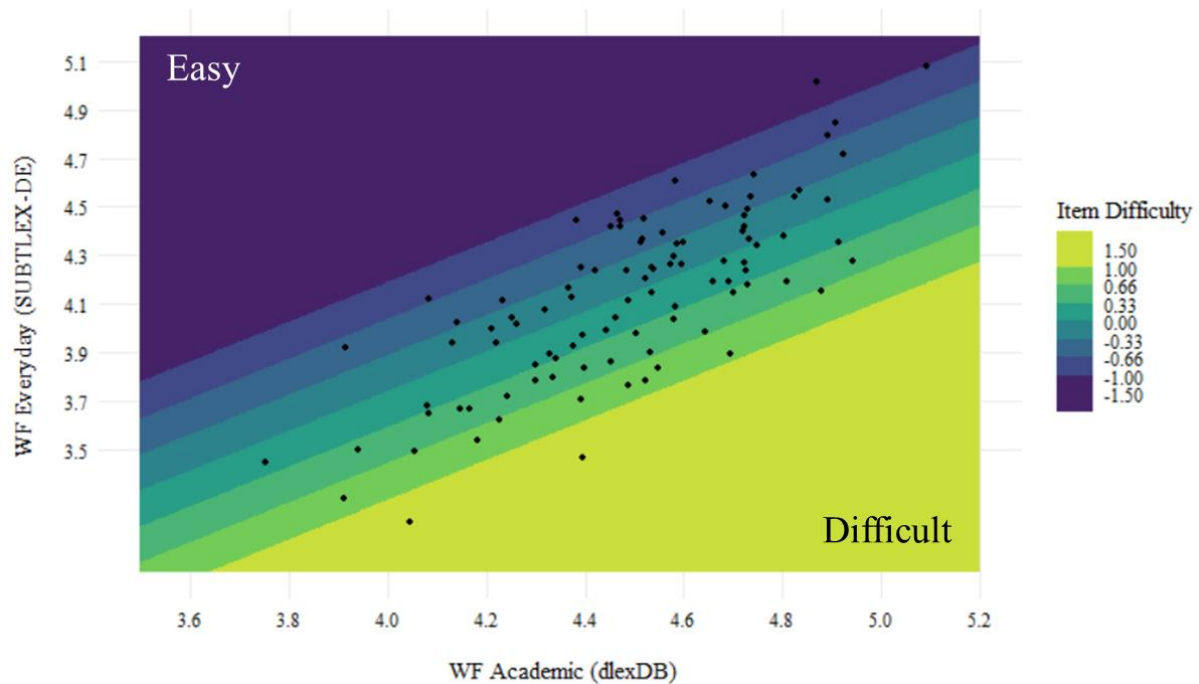
	Only Everyday			Only Academic			Both together		
	β	t	p	β	t	p	β	t	p
Dependent variable: Item difficulty									
β_1 WF Eday	-0.34	-3.51	<.001				-0.65^a	-4.50	<.001
β_2 WF Acad				-0.09	-0.01	.364	0.41^a	2.80	.006
R^2	.12			.01			.19		

Note. $N = 99$ political knowledge items, Pooled coefficient from 20 imputed datasets, $\beta =$ standardized regression coefficient. WF: Average word frequency of nouns, verbs and adjectives, bold coefficients are significant ($p < .05$). Appendix B includes the analysis with different word exclusion criteria. ^a β 's significantly different in value, $F(1, 96) = 6.59$, $p = .012$.

The regression coefficient for academic word frequency increases when everyday word frequency is entered into the model. Thus, in the combined model, academic word frequency has a statistically significant effect, $\beta_2 = 0.41$, $t = 2.81$, $p = .007$. This effect is significantly larger than the effect in the model including only academic word frequency, $SE_{ab} = 0.12$, $t = -4.26$, $p < .001$, and $a * b * c' = -0.22$. Additionally, the regression coefficient for everyday word frequency increases when academic word frequency is entered into the model, $SE_{ab} = 0.12$, $t = -2.68$, $p = .012$, $a * b * c' = -0.22$. This statistical phenomenon is known as a mutual suppression effect (MacKinnon et al., 2000) or enhancement effect (Friedman & Wall, 2005).

Figure 3

Predicted item difficulty relative to average word frequency in everyday and academic settings



Note. x-Axis: Average word frequency of nouns, verbs and adjectives from dlexDB, y-Axis: Average word frequency of nouns, verbs and adjectives in SUBTLEX-DE, points represent the *actual* bivariate distribution of items' average word frequency. The color represents the *predicted* item difficulty.

The model revealed a complimentary effect of combining word frequency in everyday and academic settings. The results suggested that the unshared variance of academic and everyday word frequency explains a substantive additional amount of variance in item difficulty. In other words, item difficulty is explained by relative frequency in each setting respectively. Items with words that are particularly frequent in everyday relative to academic settings are easier. Conversely, items with words relatively more frequent in academic settings compared to everyday settings are more difficult (see Figure 3).

These results were replicated using regression with residualized variables and the formulas suggested by Friedman and Wall (2005). These results can be found in Appendix B Tables 6 and 7, and conditions under which sign changes occur in suppressions in Appendix C Figure C1.

Additionally, we can illustrate these results with the principle components, shared frequency and everydayness (see Table 5). Items difficulty was explained by the shared frequency

component, $\beta_1 = -0.24$, $t = -2.30$, $p = .024$, and the everydayness component, $\beta_2 = -0.36$, $t = -3.78$, $p < .001$. Items were easier when words were frequent in both settings, while the relative frequency in everyday settings (i.e., Everydayness) explains additional variance.

Table 5

Regression analysis explaining item difficulty in the political knowledge test with principal components of everyday and academic word frequency

	Shared Frequency			Everydayness			Both together		
	β	t	p	β	t	p	β	t	p
Dependent variable: Item difficulty									
β_1 Shared Frequency	-0.24	-2.30	.024				-0.24^a	-2.46	.016
β_2 Everydayness				-0.36	-3.78	<.001	-0.36^a	-3.88	<.001
R^2	.05			.13			.19		

Note. $N = 99$ political knowledge items, pooled coefficient from 20 imputed datasets, $\beta =$ standardized regression coefficient. Bold coefficients are significant ($p < .05$). ^a β 's not significantly different, $F(1, 96) = 1.11$, $p = .295$.

Discussion

The present paper investigates whether word frequency could be a potential point of reference for item difficulty in political knowledge assessment. Thus, we analyzed how word frequency in everyday and academic settings explains the difficulty of 99 items in a political knowledge assessment administered to 250 German secondary school students in 7th and 10th grade. The results showed that word frequencies in everyday and academic settings significantly explain item difficulty in political knowledge assessment.

Items with words that are more frequent in academic and everyday settings are easier. Thus, we found an effect of word frequency in knowledge test items. The environmental opportunity hypothesis suggests that exposure to content and opportunities to learn plays an important role in knowledge acquisition (Stanovich & Cunningham, 1993). Word frequency could be an indicator of the likelihood of exposure to the facts and concepts addressed in the item and the language used to express them. However, based on this study, we cannot clearly attribute the effect of word frequency to the likelihood of exposure or learning opportunities. First, frequency could also be related to complexity. Inherently complex topics (e.g., European integration) are less likely to be part of informal everyday language use. Therefore, frequently occurring facts and concepts might be intrinsically easier to understand, learn, and express. Second, words with low frequency tend to be longer (Table 1: e.g., Bundesverfassungsgericht

"[Federal Constitutional Court]") and more similar to each other (Table 1: e.g., "Bundestag", "Bundesrat", "Bundesamt"). Both word length (i.e., syllable length) and similarity (i.e., OLD20) are factors that reduce readability (e.g., Fitzgerald et al., 2015). Readability is a validity issue if we assume that an item's decoding demands are higher than its knowledge demands. For instance, there could be students who know something about the Bundesverfassungsgericht [Federal Constitutional Court] but cannot decode the word "Bundesverfassungsgericht". In sum, the environmental opportunity hypothesis provides an interpretation of the word frequency effect in knowledge tests; however, further research needs to investigate the extent to which complexity, readability and other factors influence the word frequency effect in knowledge test items.

A novel finding is that combined word frequency in everyday and academic settings has a complementary effect. The relative frequency in everyday settings appears to be very important for explaining item difficulty. Items with words that are relatively frequent in everyday compared to academic settings are particularly easy. It has been argued that word frequency from corpora that best represent the language use to which individuals are actually exposed are the best predictors of difficulty in different tasks (e.g., lexical decision task; Brysbaert et al., 2019). When controlling for word frequency in everyday settings, word frequency in academic settings could become an indicator of the extent to which the item's content is disconnected from real-life experiences and everyday language use. Item contents that are closer to everyday experiences and everyday language use are easier. However, relative frequency in everyday settings could also be a better indicator for complexity and readability than academic word frequency. These interpretations need to be validated in further systematic research; nonetheless, it seems to be worthwhile to combine word frequency in different language settings when investigating item difficulty.

Overall, our results suggest that word frequency from different language settings can explain item difficulty in educational assessments. An exposure-driven approach led us to assume that word frequency would be an indicator for item difficulty; however, the mechanisms underlying the word frequency effect are probably more complex and multifaceted than merely an effect of exposure and learning opportunities. More research is needed to investigate to what extent word frequency reflects exposure probability, complexity and readability. Explaining differences in achievement between students and groups of students via differences in opportunities for learning in everyday or academic settings is an area of focus within educational psychology (e.g., Schuth et al., 2017). So there may be several applications where the frequency

of words in different contexts could help us better understand why students and groups of students perform differently.

We found that combining highly correlated variables can have a relatively large explanatory value. High correlation does not always mean redundancy (Friedman & Wall, 2005). In this case, we found a suppression effect that we consider theoretically and practically relevant. It should be noted that the results of analyses with highly correlated predictors should be interpreted with caution. Cohen et al. (2003) suggested that suppression effects could be a statistical artifact due to model instability. Contradictory to this, Friedman and Wall (2005) concluded in a study on suppression effects that "our findings indicate that multicollinearity may produce very desirable results" (p. 135). Wurm and Fisicaro (2014) concluded from a study on multicollinearity in psycholinguistic research, "[...] suppression does not indicate computational problems or model instability" (p. 47). The debate on suppression effects could be important because a recent literature review showed that one-third of the publications in psychology journals contain evidence of statistical suppression effects (Martinez Gutierrez & Cribbie, 2021). We chose to interpret our results in this manner because the OLR with a suppression effect, a regression with orthogonal principle components and several other methods (i.e., cross-validated regression, regression with residualized variables, ridge regression and Friedman and Wall's (2005) formulas) yielded consistent results.

Strength & Limitations

In this paper, we did not build a comprehensive model of item difficulty in political knowledge tests, nor did we examine an exhaustive set of item features. However, word frequency explained 19% of the variance in item difficulty. Word frequencies are very objective, reproducible, and labor-efficient variables and could help to improve item construction. Nonetheless, there is no doubt that other features (e.g., distractor plausibility, structural characteristics of political knowledge, etc.) can potentially explain additional variance in item difficulty.

Word frequency is an indicator for exposure and often used as a measure of linguistic complexity. On the one hand, conveying niche political facts and concepts will naturally require the use of words with rare frequencies. On the other hand, item difficulty could be influenced by inauthentic and inappropriate use of rare words, making the items overly complex. Our interpretation rests on the assumption that the language in knowledge test items is appropriate. Unfortunately, word frequency does not separate appropriate from inappropriate language use. Appropriateness and authenticity are matters for human judgment. The items were constructed based on international item construction guidelines (Gierl et al., 2017) and reviewed multiple

times by experts independently with the objective of ensuring an appropriate level of language complexity. Thus, we have reason to believe that the items use appropriate and authentic language.

The corpora used to determine word frequency in academic and in everyday settings stem from different modalities and are relatively old. The word frequency in everyday settings better captures language exposure via listening, while the word frequency in academic settings better captures language exposure via reading. This modality shift is not generally in conflict with the language setting because everyday language is mostly cultivated through oral communication and academic language is more frequently cultivated through written language. However, a corpus of academic language in oral settings would allow for the more straightforward interpretation of our results. In addition, most of the language sources in both corpora are older than 10 years, and students between the ages of 11 and 18 are unlikely to have consumed them when they were published. Unfortunately, we do not know of corpus in German that would be more suitable for this analysis.

The number of participants was relatively small compared to other studies in educational assessment. However, Rasch models are usually applicable in studies with smaller sample sizes (e.g., Stone & Yumoto, 2004). Additionally, we used a plausible value procedure to account for methodological issues caused by uncertainty in point estimates (e.g., Marsman et al., 2016). Nonetheless, the results should be replicated in future studies.

The study is based on a non-representative sample of students. However, a core assumption in item response theory is that unbiased estimates of item properties can be obtained from unrepresentative samples (Embretson & Reise, 2013). Therefore, the results concerning item difficulty should be largely reproducible in representative samples. Nonetheless, possible differential item functioning (e.g., Holland & Wainer, 2012) between students from different family backgrounds, for instance, in items with more everyday and academic content would undoubtedly be a very interesting research topic. However, this study's statistical power is too limited to detect such effects. Therefore, this is another aspect that should be investigated in future research.

Conclusion

People know little about the things they have little to do with. What is striking, however, is the fact that word frequencies seem to provide a simple way to describe exposure to and academic orientation of an item. Thus, word frequencies may be a fruitful indicator to improve our understanding of educational assessment in different language-related domains, not just language

testing. We publish our analysis scripts, including item parsing, multiple imputation of non-found types and plausible value drawing. SUBTLEX and corpora similar to dlexDB exist for many different languages. We encourage other researchers to replicate the analysis for different languages and knowledge domains.

References

- Alscher, P., Ludewig, U., & McElvany, N. (2022). Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung. *Zeitschrift für Erziehungswissenschaft*, 1-21. <https://doi.org/10.1007/s11618-022-01085-0>
- Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Psychology press. <https://doi.org/10.4324/9780203774441>
- Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis. *Open Journal of Statistics*, 5(07), 754. <http://dx.doi.org/10.4236/ojs.2015.57075>
- APA (2020). APA GUIDELINES for Psychological Assessment and Evaluation. <https://www.apa.org/about/policy/guidelines-psychological-assessment-evaluation.pdf>
- Benoit, K., Muhr, D., & Watanabe, K. (2021). stopwords: Multilingual Stopword Lists. R package version 2.3. <https://CRAN.R-project.org/package=stopwords>
- Friedman, L., & Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple linear regression. *The American Statistician*, 59(2), 127-136. <https://doi.org/10.1198/000313005X41337>
- Berendes, K., Vajjala, S., Meurers, D., Bryant, D., Wagner, W., Chinkina, M., & Trautwein, U. (2018). Reading demands in secondary school: Does the linguistic complexity of textbooks increase with grade level and the academic orientation of the school track?. *Journal of Educational Psychology*, 110(4), 518-543. <https://doi.org/10.1037/edu0000225>
- Bischof, D., & Senninger, R. (2018). Simple politics for the people? Complexity in campaign messages and political knowledge. *European Journal of Political Research*, 57(2), 473-495. <https://doi.org/10.1111/1475-6765.12235>
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental Psychology*, 58, 412-424. <https://doi.org/10.1027/1618-3169/a000123>
- Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27(1), 45-50. <https://doi.org/10.1177/0963721417727521>
- Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, 51(2), 467-479. <https://doi.org/10.3758/s13428-018-1077-9>

- Care, E., Griffin, P., & Wilson, M. (2018). *Assessment and Teaching of 21st Century Skills*. Springer.
- Carpini, M. D., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press. <http://www.jstor.org/stable/10.2307/j.ctt1cc2kv1>
- Chen, X., & Meurers, D. (2018). Word frequency and readability: Predicting the text-level readability with a lexical-level attribute. *Journal of Research in Reading*, *41*(3), 486-510. <https://doi.org/10.1111/1467-9817.12121>
- Clark, N. (2017). Explaining political knowledge: the role of procedural quality in an informed Citizenry. *Political Studies*, *65*(1), 61-80. <https://doi.org/10.1177/0032321716632258>
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, *34*(2), 213-238. <https://doi.org/10.2307/3587951>
- Cramer, K. J., & Toff, B. (2017). The fact of experience: Rethinking political knowledge and civic competence. *Perspectives on Politics*, *15*, 754-770. <https://doi.org/10.1017/S1537592717000949>
- Darlington, R. B. (1990). *Regression and linear models*. Guilford Press.
- De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational psychologist*, *31*(2), 105-113. https://doi.org/10.1207/s15326985ep3102_2
- Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, *66*(5), 843-863. <https://doi.org/10.1080/17470218.2012.720994>
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press. <https://doi.org/10.4324/9781410605269>
- Embretson, S., & Yang, X. (2006). Automatic Item Generation and Cognitive Psychology. *Handbook of statistics*, *26*, 747-768. [https://doi.org/10.1016/S0169-7161\(06\)26023-1](https://doi.org/10.1016/S0169-7161(06)26023-1)
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191. <https://doi.org/10.3758/BF03193146>
- Fitzgerald, J., Elmore, J., Koons, H., Hiebert, E. H., Bowen, K., Sanford-Moore, E. E., & Stenner, A. J. (2015). Important text characteristics for early-grades text complexity. *Journal of Educational Psychology*, *107*(1), 4-29. <https://doi.org/10.1037/a0037289>
- Frey, A., Hartig, J., & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational*

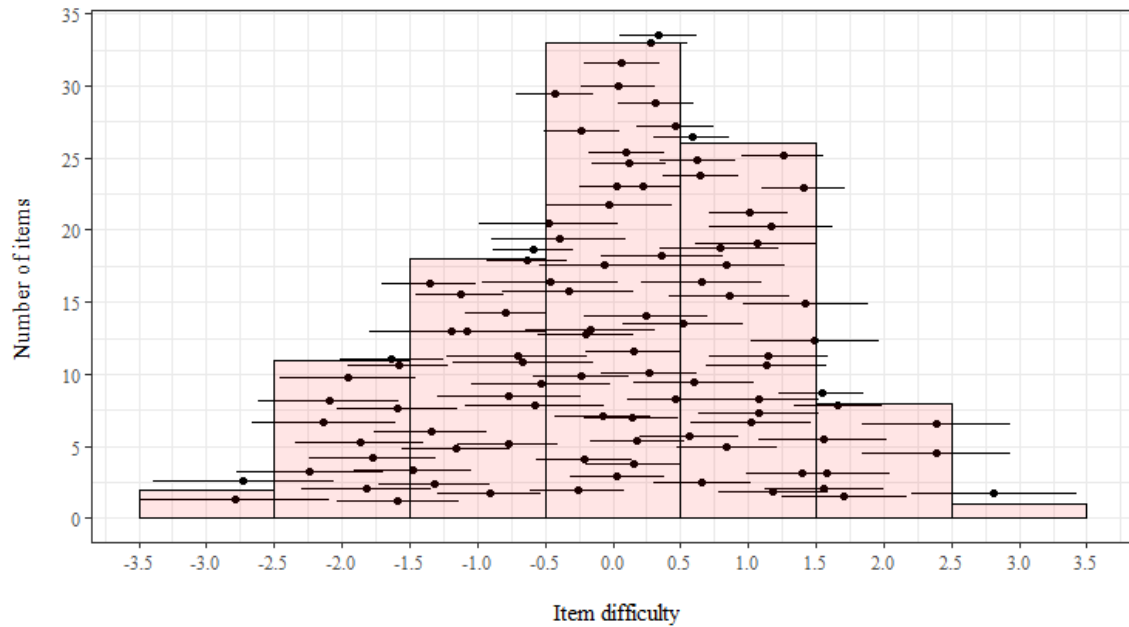
- measurement: issues and practice*, 28(3), 39-53. <https://doi.org/10.1111/j.1745-3992.2009.00154.x>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1), 1-22. <https://www.jstatsoft.org/v33/i01/>
- Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: a comprehensive review. *Review of Educational Research*, 87(6), 1082-1116. <https://doi.org/10.3102/0034654317726529>
- Holland, P. W., & Wainer, H. (2012). *Differential Item Functioning*. Routledge. <https://doi.org/10.4324/9780203357811>
- Irwing, P., Cammock, T., & Lynn, R. (2001). Some evidence for the existence of a general factor of semantic memory and its components. *Personality and Individual Differences*, 30(5), 857-871. [https://doi.org/10.1016/S0191-8869\(00\)00078-7](https://doi.org/10.1016/S0191-8869(00)00078-7)
- Juhasz, B. J., Yap, M. J., Raoul, A., & Kaye, M. (2019). A further examination of word frequency and age-of-acquisition effects in English lexical decision task performance: The role of frequency trajectory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(1), 82–96. <https://doi.org/10.1037/xlm0000564>
- Kim, J. S., Burkhauser, M. A., Mesite, L. M., Asher, C. A., Relyea, J. E., Fitzgerald, J., & Elmore, J. (2021). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology*, 113(1), 3–26. <https://doi.org/10.1037/edu0000465>
- Krathwohl, D. R., & Anderson, L. W. (2010). Merlin C. Wittrock and the revision of Bloom's taxonomy. *Educational Psychologist*, 45(1), 64-65. <https://doi.org/10.1080/00461520903433562>
- Kuhn, M. (2021). caret: Classification and Regression Training. R package version 6.0-90. <https://CRAN.R-project.org/package=caret>
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, 1(4), 173-181. <https://doi.org/10.1023/A:1026595011371>
- Marsman, M., Maris, G., Bechger, T., & Glas, C. (2016). What can we learn from Plausible Values?. *Psychometrika*, 81(2), 274-289. <https://doi.org/10.1007/s11336-016-9497-x>
- Martinez Gutierrez, N., & Cribbie, R. (2021). Incidence and interpretation of statistical suppression in psychological research. *Canadian Journal of Behavioural Science*, 53(4), 480–488. <https://doi.org/10.1037/cbs0000267>

- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3-62. https://doi.org/10.1207/S15366359MEA0101_02
- Mustillo, S., & Kwon, S. (2015). Auxiliary variables in multiple imputation when data are missing not at random. *The Journal of Mathematical Sociology*, 39(2), 73-91.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91-108. <https://doi.org/10.1002/RRQ.011>
- National Research Council. (2012). A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas. *National Academies Press*.
- Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. S. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods*, 40(4), 1001-1015. <https://doi.org/10.3758/BRM.40.4.100>
- R Core Team (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. <http://www.R-project.org/>
- Reif, F., & Heller, J. I. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist*, 17(2), 102-127. <http://dx.doi.org/10.1080/00461528209529248>
- Robitzsch, A., Kiefer, T., & Wu, M. (2018). TAM: Test analysis modules. *R package version, 2*, 9-35. <https://github.com/alexanderrobitzsch/TAM>
- Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction*, 49, 157-165.
- Segedinac, M. T., Horvat, S., Rodić, D. D., Rončević, T. N., & Savić, G. (2018). Using knowledge space theory to compare expected and real knowledge spaces in learning stoichiometry. *Chemistry Education Research and Practice*, 19(3), 670-680. <https://doi.org/10.1039/C8RP00052B>
- Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods*, 6(4), 317.
- Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, 85(2), 211-229. <https://doi.org/10.1037/0022-0663.85.2.211>
- Stefanutti, L., Heller, J., Anselmi, P., & Robusto, E. (2012). Assessing the local identifiability of probabilistic knowledge structures. *Behavior Research Methods*, 44(4), 1197-1211. <https://doi.org/10.3758/s13428-012-0187-z>

- Stevenson, C. E., Hickendorff, M., Resing, W. C., Heiser, W. J., & de Boeck, P. A. (2013). Explanatory item response modeling of children's change on a dynamic test of analogical reasoning. *Intelligence*, *41*(3), 157-168. <https://doi.org/10.1016/j.intell.2013.01.003>
- Stone, M., & Yumoto, F. (2004). The effect of sample size for estimating Rasch/IRT parameters with dichotomous items. *Journal of Applied Measurement*, *5*(1), 48–61.
- Taikh, A., Hargreaves, I. S., Yap, M. J., & Pexman, P. M. (2015). Semantic classification of pictures and words. *The Quarterly Journal of Experimental Psychology*, *68*(8), 1502–1518. <https://doi.org/10.1080/17470218.2014.975728>
- Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 453–488). Lawrence Erlbaum Associates, Inc. <https://apps.dtic.mil/sti/pdfs/ADA183189.pdf>
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavior Research Methods*, *45*(4), 1115-1143. <https://doi.org/10.3758/s13428-012-0307-9>
- Vajjala, S., & Meurers, D. (2012). *On improving the accuracy of readability classification using insights from second language acquisition*. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*. Association of Computational Linguistics. <https://aclanthology.org/W12-2019.pdf>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1-67. <https://www.jstatsoft.org/v45/i03/>
- Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190. <https://doi.org/10.1080/17470218.2013.850521>
- Weißeno, G., Detjen, J., Juchler, I., Massing, P., & Richter, D. (2010). *Konzepte der Politik – ein Kompetenzmodell* [Concepts of politics - A competence model]. <http://nbn-resolving.org/urn:nbn:de:0111-pedocs-120091>
- Westle, B., & Tausendpfund, M. (2019). Politisches Wissen: Relevanz, Messung und Befunde [Political knowledge: Relevance, Measurement and Findings]. In *Politisches Wissen [Political knowledge]* (pp. 1-39). Springer. https://doi.org/10.1007/978-3-658-23787-5_1

- Wise, S. L., & DeMars, C. E. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurement*, 43(1), 19-38.
<https://psycnet.apa.org/doi/10.1111/j.1745-3984.2006.00002.x>
- Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, 72, 37-48.
<https://doi.org/10.1016/j.jml.2013.12.003>

Appendix A

Figure A1*Distribution of item difficulty*

Note. X-axis scale of item difficulty ranging from -3.5 to 3.5 on a logit scale. Points represent mean difficulty estimate for each item. The range around the points represents the 95% confidence interval of the item difficulty estimate. The y-axis represents the number of items within a binwidth of 1. Point position is based on a frequency count within the binwidth and random variation (jitter) to avoid overlapping confidence intervals.

Appendix B

Table B1

Correlations and descriptive statistics

Variables	1	2	3	4	5	6	7
WF everyday:							
1. All words							
2. Without stop words	.87						
3. Nouns, verbs, and adjectives	.86	.93					
WF academic							
4. All words	.82	.67	.67				
5. Without stop words	.57	.75	.70	.77			
6. Nouns, verbs, and adjectives	.60	.70	.77	.79	.93		
7. Item difficulty ^a	-.29	-.37	-.34	-.06	-.15	-.09	
<i>M</i>	5.14	4.14	4.15	5.35	4.49	4.49	0
<i>SD</i>	0.27	0.33	0.35	0.18	0.24	0.26	1.18

Note. $N = 99$ Political knowledge items. Bold correlations, $p < .001$. WF: Average word frequency. Dark grey field: Correlations between everyday and academic word frequency with the same criteria. Light grey fields: Correlations with item difficulty. ^aPlease find histogram displaying the distribution of item difficulty in Appendix A.

Table B2

Regression analysis explaining item difficulty in the political knowledge test with different word exclusion criteria

	Only Everyday			Only Academic			Both together		
	β	<i>t</i>	<i>p</i>	β	<i>t</i>	<i>p</i>	β	<i>t</i>	<i>p</i>
All words									
β_1 WF Eday	-0.29	-2.91	.005				-0.72	-4.41	<.001
β_2 WF Acad				-0.06	-0.58	.567	0.53	3.23	.002
R^2	.08			<.01			.18		
Without stopwords									
β_1 WF Eday	-0.37	-3.85	<.001				-0.59	-4.17	<.001
β_2 WF Acad				-0.15	-1.43	.157	0.29	2.08	.041
R^2	.14			.02			.18		
Nouns, verbs and adjectives									
β_1 WF Eday	-0.36	-3.71	<.001				-0.68	-4.54	<.001
β_2 WF Acad				-0.12	-1.13	.263	0.41	2.74	.007
R^2	.12			.01			.19		

Note. $N = 99$ political knowledge items, Pooled coefficient from ten imputed datasets, β = standardized regression coefficient. WF: Average word frequency. Eday: Everyday, Acad: Academic. Bold coefficients are significant ($p < .05$).

Appendix C

Table C1

Regression analysis explaining item difficulty in the political knowledge test with residualized academic word frequency variables

	Only WF Everyday			Only $e(\text{Acad} \rightarrow \text{Eday})$			Both together		
	β	t	p	β	t	p	β	t	p
Dependent variable: Item difficulty									
β_1 WF Eday	-0.34	-3.51	<.001				-0.34^a	-3.63	<.001
β_2 $e(\text{Acad} \rightarrow \text{Eday})$				0.26	2.64	.010	0.26^a	2.81	.006
R^2	.12			.07			.19		

Note. $N = 99$ political knowledge items, pooled coefficient from 20 imputed datasets, β = standardized regression coefficient. WF: Average word frequency of nouns, verbs and adjectives. Eday: Everyday, Acad: Academic. Bold coefficients are significant ($p < .05$). ^a β 's significantly different, $F(1, 96) = 22.68, p < .001$.

Table C2

Regression analysis explaining item difficulty in the political knowledge test with residualized everyday word frequency variables

	Only WF Academic			Only $e(\text{Eday} \rightarrow \text{Acad})$			Both together		
	β	t	p	β	t	p	β	t	p
Dependent variable: Item difficulty									
β_1 WF Acad	-0.09	-0.91	.364				-0.09 ^a	-1.00	.321
β_2 $e(\text{Eday} \rightarrow \text{Acad})$				-0.42	-4.49	<.001	-0.42^a	-4.49	<.001
R^2	.01			.18			.19		

Note. $N = 99$ political knowledge items, pooled coefficient from 20 imputed datasets, β = standardized regression coefficient. WF: Average word frequency of nouns, verbs and adjectives. Eday: Everyday, Acad: Academic. Bold coefficients are significant ($p < .05$). ^a β 's significantly different, $F(1, 96) = 5.48, p < .011$.

Table C3

Computation of β 's and R^2 given the correlation (Friedman & Wall, 2005) between everyday word frequency and item difficulty ($r_{y1} = -.34$), academic word frequency and item difficulty ($r_{y2} = -.09$), and between everyday and academic word frequency ($r_{12} = .77$)

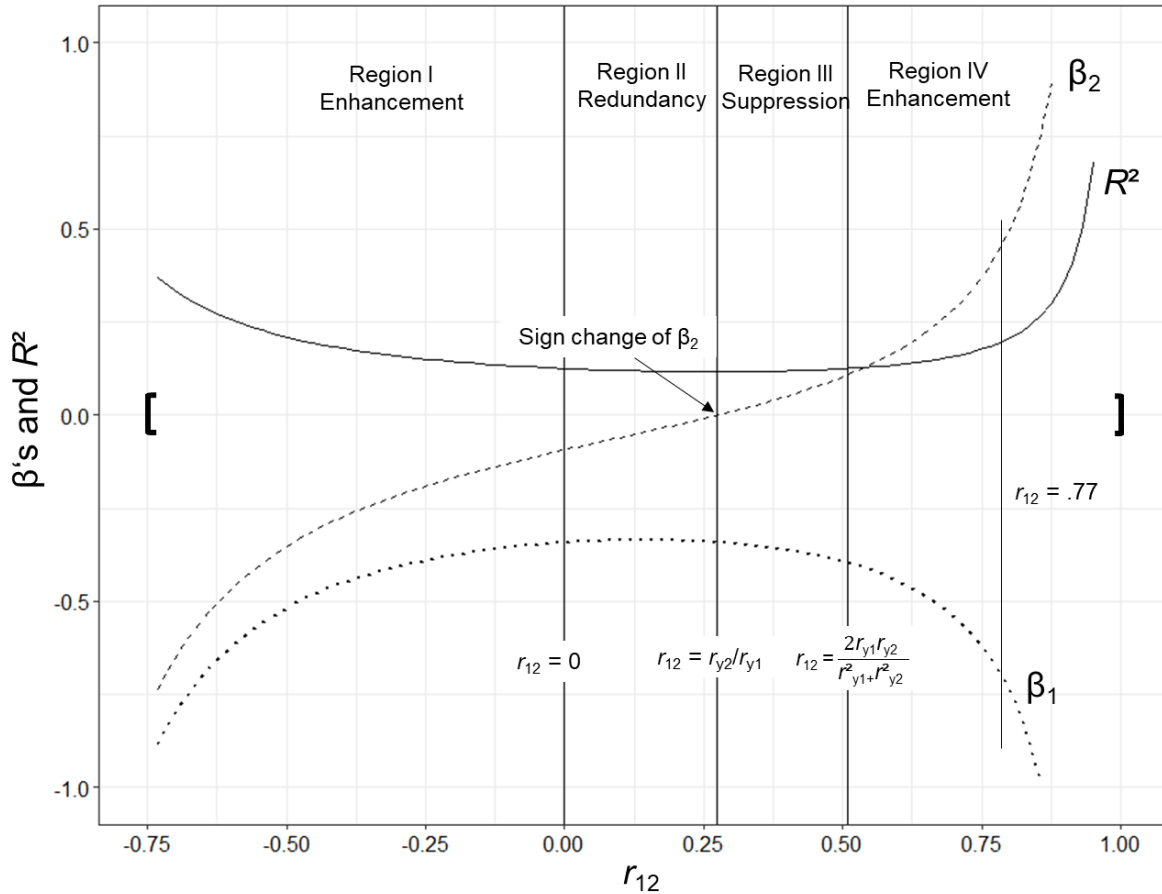
1	$F1: \beta_1 = \frac{r_{y1} - r_{y2} * r_{12}}{1 - r_{12}^2} = -0.65$
2	$F2: \beta_2 = \frac{r_{y2} - r_{y1} * r_{12}}{1 - r_{12}^2} = 0.41$
3	$F3: R^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2r_{y1}r_{y2}r_{12}}{1 - r_{12}^2} = .19$
	“[...] there are no limits on multicollinearity in regression on two predictors other than those given by the necessity to have a nonnegative definite matrix” (p.135) Matrix is nonnegative definite if:
4	$F4: r_{y1} * r_{y2} - \sqrt{(1 - r_{y1}^2)(1 - r_{y2}^2)} \leq r_{12} \leq r_{y1} * r_{y2} + \sqrt{(1 - r_{y1}^2)(1 - r_{y2}^2)} =$ $= -.90 \leq .77 \leq .97$
5	“We suggest that the Regions I-IV, delineated in this article, provide a clear structure by which the correlations involved in a regression on two predictors can be analyzed.”

(See Appendix C Figure 1)

Note. Exact estimates for correlations with only nouns, verbs and adjectives: $r_{y1} = -.3415696$, $r_{y2} = -.09357449$, $r_{12} = .7666544$.

Figure C1

Graphical display of β 's and R^2 given the correlation between everyday word frequency and item difficulty ($r_{y1} = -.34$), academic word frequency and item difficulty ($r_{y2} = -.09$), as well as all possible values for the correlation between everyday and academic word frequency (r_{12})



Note. β 's and R^2 for $r_{y1} = -.34$, $r_{y2} = -.09$ and all possible values of r_{12} . The sign change (i.e., negative correlation r_{y2} results in positive β_2) occurs under the conditions $r_{12} > r_{y2}/r_{y1}$ ($r_{y2}/r_{y1} = .33$). Given the correlations $r_{y1} = -.34$ and $r_{y2} = -.09$, a sign change will occur if the word frequencies are correlated $r_{12} > .26$. Exact estimates for correlations with nouns, verbs and adjectives: $r_{y1} = -.3415696$, $r_{y2} = -.09357449$, $r_{12} = .7666544$.

2.4 Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators

Alscher, P., Ludewig, U., & McElvany, N. (2023). Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators. *Journal of Youth and Adolescence*, *51*, 1886–1900. <https://doi.org/10.1007/s10964-022-01639-9>

Abstract

Civic education is generally assumed to play a key role in youth's political sophistication. It aims to equip young people with the necessary competencies and skills to effectively participate in political and civic life. However, few studies have examined the relative importance of different facets of teaching quality within civic education as well as mediating factors for fostering active citizens. The present study seeks to fill this gap by investigating how different facets of teaching quality are associated with adolescents' willingness to participate in political and civic life and how this relationship is mediated by political knowledge and interest. The study uses original data from $N = 250$ students ($n = 152$ 7th graders: $M_{age} = 12.54$, $SD = 0.91$, range = 11–14, 45% female; $n = 98$ 10th graders: $M_{age} = 16.12$, $SD = 0.97$, range = 15–18, 35% female). The findings show that not all teaching quality facets are equally important. While perceived cognitive activation and open classroom climate were positively associated with students' willingness to participate, a statistically significant association with discussions of current political events in the classroom was not found. In addition, the relationship between perceived cognitive activation and willingness to participate is fully mediated by students' political knowledge and interest. This study illustrates the relative importance of different teaching quality facets in civic education and calls for continued efforts to better understand teaching quality in civic education.

Keywords: Civic Education · Cognitive Activation · Open Classroom Climate · Political Participation · Adolescence

Introduction

Democratic institutions and laws are hardly sufficient to ensure the survival of democracy (e.g., Almond & Verba, 1963; Putnam, 2000). Instead, democracies and their institutions can only function well if embedded in a *culture* of democracy (see Inglehart & Welzel, 2003). Accordingly, the quality and persistence of democracy depends on the participation and underlying political sophistication of a country's citizenry (see Milner, 2002). In general, political sophistication can refer to an individual's political knowledge and interest. When looking at conventional forms of participation like political party engagement or voting, one can usually observe a participation gap between younger and older generations (see Maurissen, 2020). This is accompanied by worries that youth's interests will be underrepresented in politics if younger generations participate systematically less than older generations. Such a development might threaten democracy in the long run (see Barrett, 2018). A major cornerstone within the process of political sophistication and a crucial resource for encouraging youth to participate in political and civic life is formal, school-based civic education (see Galston, 2001). Various aspects of teaching quality play an important role in civic education, offering teachers ways to engage students with political and civic content, encourage them to build political knowledge and interest, and ultimately equip them with the necessary competencies and skills to effectively participate in political and civic life. Despite the accumulating body of evidence in this field, crucial research gaps persist. Little is known about which teaching quality facets are of greatest importance in civic education. This study aims to link teaching quality through interest and knowledge with students' willingness to participate in political and civic life. It investigates the relative importance of different civic education teaching quality facets for students' willingness to participate and the role that political interest and political knowledge play in this relationship.

Civic Education in Schools

Schools have a civic mission, which is to produce an informed and engaged citizenry (Campaign for the Civic Mission of Schools, 2011). School-based civic education plays an important role in fulfilling this mission (Kahne et al., 2013). It aims to equip students with the knowledge, democratic skills and attitudes necessary to effectively participate in political and civic life. It can be conceptualized as all opportunities provided by schools to engage their students in meaningful learning experiences to facilitate and enhance their development as responsible and active citizens (Homana et al., 2006). Alongside national-level and student-level aspects (see

Isac et al., 2011), class-level variables like the quantity and quality of civic learning opportunities are particularly important for the success of civic education (e.g., Hess & McAvoy, 2015).

Political and Civic Participation as a Major Outcome of Civic Education

In civic education research, it is widely accepted that political and civic participation is an integral part of a well-functioning democratic system and therefore a desired outcome of civic education. In addition to being a key defining element of democracy, political and civic participation is an expression of adolescents' healthy and positive development (see Manganelli et al., 2015). In Germany, the standing conference of the ministers of education and cultural Affairs (*Kultusministerkonferenz*) claims that encouraging students to stand up for freedom, democracy, human rights etc. must be a vital objective of school education (Kultusministerkonferenz, 2018). Besides, attitudes towards political and civic participation usually manifest during adolescence (Baumert et al., 2016), reinforcing its position as an important outcome of civic education in schools.

The majority of political and civic participation research among adolescents is based on willingness to participate rather than actual participation. This is primarily due to adolescents' limited opportunities and other constraints that both directly (e.g., lack of voting rights) and indirectly (e.g., lack of transportation, family responsibilities) affect their participation (McWhirter & McWhirther, 2016). Looking at intentions rather than actual behavior is based on the theory of planned behavior (Ajzen, 1991), which postulates that intentions are immediate antecedents of actual behavior. Indeed, empirical evidence from the U.S. and Italy suggests that voting intentions and voting behavior have a strong positive relationship (Ajzen et al., 1982; Roccato & Zogmaister, 2010). Willingness to participate is therefore considered the closest available proxy indicator of future political and civic participation (see Quintelier & Hooghe, 2013).

Data from the Shell Youth Studies have shown that school type matters for students' voting intentions. More precisely, German 10th graders attending schools with an academic track have greater voting intentions than students attending schools with a vocational track (Wallrich et al., 2021). Similarly, educational attainment has a positive effect on adolescents' voting intentions (Yang and Hoskins, 2020) and young adults at age 20 have higher self-reported actual voting levels if they completed an academic track in school compared to a vocational track (Janmaat et al., 2014).

In Germany, students are separated into different secondary tracks at the end of elementary school, which usually happens at around age 10. The school assignment procedure is

primarily based on students' achievements and abilities. Although there is considerable variation across German states, the most common system is the "tripartite" system of Hauptschule, Realschule and Gymnasium. Among the three, Hauptschule is the most vocationally oriented track. Realschule is the intermediate track and Gymnasium is the academically most challenging track. For clarity, the study differentiates between schools that prepare for vocational education and training (i.e. Hauptschule and Realschule) and Gymnasium, which prepares students for university entrance. Usually, the proportion of students from immigrant backgrounds is higher and average socio-economic status lower among students in schools preparing for the vocational track (Wallrich et al., 2021). In addition, research has shown that students in the vocational track have less access to civic education classes (Achour & Wagner, 2020).

Furthermore, research suggests that adolescents' willingness to participate in political and civic life depends on their socio-economic background. For instance, in the Italian subsample of the ICCS 2009 data, the national index of socio-economic background, comprising parents' highest occupational status, parents' highest educational level, and the number of books at home, is significantly positively associated with students' expected electoral expectation (Manganelli et al., 2012). This result was replicated when considering political knowledge and efficacy beliefs as mediating variables (Manganelli et al., 2014). However, empirical evidence suggests that effective civic education through high teaching quality can weaken the link between socio-economic background and willingness to participate (e.g., Castillo et al., 2015).

Despite some national differences, most countries' education systems share a rather broad concept of citizenship (Claes & Hooghe, 2009). In Western European countries, teacher beliefs regarding the aims of civic education are very similar (Reichert & Torney-Purta, 2019). In addition, ICCS 2016 data show that the relationship between civic education and civic outcomes is rather stable across countries. In all 24 participating countries, political interest and knowledge were significantly positively associated with students' willingness to participate in.

Teaching Quality in Civic Education

Teaching quality has a long tradition in educational research and has proven to be a relevant feature of schools' effectiveness (see Hattie, 2012). In general, teachers play a crucial role in determining how civic education is taught. Not only are teachers able to adjust the content of civic education classes within the outlines of the official curriculum to students' needs, they can also adjust their classroom behavior.

From the perspective of deliberative democracy theory, which stresses the importance of discussion and deliberation in democracy (see Goodin & Spiekermann, 2018), an important

facet of teaching quality in civic education is an open classroom climate – that is, a classroom climate in which controversial issues are openly discussed with respect for everyone’s opinion (Godfrey & Grayman, 2014). A multitude of studies have shown that adolescents in classes with frequent and open deliberation perform better on various civic outcomes (e.g., Campbell, 2008; Eckstein et al., 2021; Manganeli et al., 2015). Data from the International Citizenship and Education Study (ICCS) 2016 revealed that in all participating countries, students’ perceptions of an open classroom climate were on average positively associated with students’ political knowledge and political interest (Schulz, Ainley, et al., 2018). Similar findings were reported for predecessor studies (Schulz et al., 2009; Torney-Purta et al., 2001).

Another relevant facet of teaching quality in civic education is the discussion of current political events in the classroom. Deliberating about current political controversies and events provides opportunities for students to consider diverse perspectives and develop political opinions (McCafferty-Wright & Knowles, 2016). Furthermore, it equips students with skills to engage with the political discourse and democratic life beyond school (Hess & McAvoy, 2015) and enables them to gain political knowledge apart from their textbooks. In particular, exposing students to current political controversies might spark their political interest. Based on data from the 2010 NAEP 12th grade Civics Assessment, it could be shown that while controlling for students’ demographic characteristics and aspects of their home environment, talking about current political events at any time in class was positively associated with scores on the NAEP Civics Assessment (Bittman and Russel, 2016). Moreover, data from the California Survey of Civic Education, comprising 2,366 California high school seniors shows that 61% of students who reported frequent talk about current political events in class said that they were interested in politics, compared to only 32% of students who reported no talk about current political events in class (Kahne, 2005).

Another facet of teaching quality in civic education that has been rather overlooked so far is cognitive activation. Cognitive activation describes teachers’ abilities to create learning activities with a greater potential for higher-level thinking, such as the provision of appropriately challenging teaching material, capitalization on students’ prior knowledge and grappling with the learning content in an in-depth way (e.g., Kunter & Trautwein, 2018; Stang & McElvany, 2020). In civic education particularly, cognitive activation could refer to encouraging students to shift their perspective and review political decisions regarding their argumentative robustness. Furthermore, cognitively activating tasks equip students with collective problem solving skills. Cognitive activation has proven to be an effective tool for stimulating students’ learning and interest in other subjects (e.g., mathematics: Baumert et al., 2010; science learning:

Fauth et al., 2014; Biology: Förtsch et al., 2016), while empirical evidence about its potential role in civic education is limited. But some initial results from Germany suggest that perceived cognitive activation has a positive effect on 9th and 10th graders' civic education-specific content knowledge (Weißeno and Landwehr, 2015).

The exact relationship between civic education teaching quality and willingness to participate has not yet been well researched. Theoretically, two processes are conceivable as to how teaching quality is related to willingness to participate. The first possibility is that teaching quality is directly related to willingness to participate. Theoretical considerations regarding a direct relationship are scarce. One possible explanation could be that increased teaching quality leads to increased expectations regarding one's own behaviors. However, there seems to be no empirical evidence to support such a direct relationship. The second possibility is that civic education teaching quality is only indirectly related to willingness to participate. The relationship between civic education and willingness to participate might be mediated by more immediate goals of civic education, such as political knowledge and political interest (see Bayram-Özdemir et al., 2016, Maurissen, 2020).

Political Knowledge and Political Interest as Mediators

Raising students' political knowledge and stimulating their political interest are essential immediate goals of civic education instruction (Claes & Hooghe, 2008). At the same time, students' political knowledge and political interest are of crucial importance in determining voting habits (e.g., Gil de Zúñiga & Diehl, 2019; Grobshäuser & Weißeno, 2020) and predicting other political and civic activities (e.g., Quintelier & Hooghe, 2012; Reichert, 2016). Therefore, civic education teaching quality may be indirectly associated with students' willingness to participate via political knowledge and political interest (Maurissen, 2020). A distinction between different immediacies of civic education goals is similarly reflected in moral development and character education theory. On the one hand, traditional or direct citizenship education explicitly focuses on knowledge transfer, which can facilitate the subsequent acquisition of other political and participation skills. On the other hand, indirect civic education places less emphasis on knowledge acquisition and focuses on moral reasoning and democratic values, thus attitudes and intentions (Clase & Hooghe, 2009). So far, these considerations have remained limited to the question of what citizenship norms civic education should pursue. This study expands their application by integrating the two approaches and defining more immediate (i.e. political interest and knowledge) and less immediate (i.e. willingness to participate) civic education outcomes.

Political knowledge involves “the range of factual information about politics that is stored in long-term memory” (Carpini & Keeter, 1996, p. 10). It is a crucial immediate goal of civic education and it relates to political and civic participation. Usually, students’ knowledge gains are evaluated with standardized tests, which enable teachers to directly evaluate students’ progress. Furthermore, political knowledge has repeatedly proved to be an important prerequisite for political and civic participation (e.g., Grobshäuser & Weißeno, 2020). Even a little knowledge about politics can empower citizens to make meaningful and informed decisions by virtue of judgment heuristics (Niemi & Junn, 1998).

Political interest is another important and immediate goal of civic education and a prerequisite for political and civic participation. Empirical evidence suggests that motivation for learning based on interest stimulates the learning process and positively affects learning outcomes (Krapp, 2002). Accordingly, stimulating students’ (domain-specific) interest is of crucial importance for scholastic achievement. Many studies suggest that political interest usually manifests during youth and changes little afterwards (Baumert et al., 2016). This further stresses the role of schools and civic education in particular in the process of stimulating youth’s political interest. Moreover, it is a well-established notion that political interest is an effective prerequisite for meaningful political and civic participation as well as for gathering political information (e.g., Claes & Hooghe, 2017; Prior & Bougher, 2018).

Current Study

Enabling young people’s effective participation in political and civic life in modern-day democracies must be an essential objective of education, and of civic education in particular. Despite the growing amount of research in the field, gaps still exist concerning the relative importance of different facets of teaching quality and the relationship between civic education and students’ participation in political and civic life. First, this study examines how perceptions of (i) an open classroom climate, (ii) discussions of current political events, and (iii) cognitive activation within civic education are associated with students’ willingness to participate in political and civic life. It is hypothesized that all three facets are positively related to willingness to participate even when considered simultaneously. In addition, the relative importance of each facet will be explored. Second, this study investigates the question of how the relationship between these perceived teaching quality facets and willingness to participate is mediated by (i) political knowledge and (ii) political interest. It is hypothesized that the relationships between the perceived teaching quality facets and students’ willingness to participate are mediated by political knowledge and political interest.

Methods

Participants

The data for this research project was collected in fall 2020 in the Study on the Development of Political and Civic Competence in Adolescence (German: *Studie zur Entwicklung politischer und gesellschaftlicher Kompetenz im Jugendalter, EPKO*). The participating students were attending 7th or 10th grade. In some German federal states, students enter secondary education in 7th grade, whereas in most others states, secondary education begins in 5th grade. In addition, students in the vocational track leave school after the 10th grade and typically enter vocational training. The period from 7th to 10th grade thus marks the core of secondary education in Germany. The schools were located in the two federal states of North Rhine-Westphalia and Thuringia. A total of 17 classes with $N = 250$ students ($n = 152$ 7th graders: $M_{age} = 12.54$, $SD = 0.91$, range = 11–14, 45% female; $n = 98$ 10th graders: $M_{age} = 16.12$, $SD = 0.97$, range = 15–18, 35% female) participated in the study. The response rate was 73%. Students not reached either did not have informed consent or were absent on the test day due to illness or quarantine measures. According to the curricula, almost all participating students had already received civic education classes as part of so-called contingent lessons (*Kontingenzstunden*), which allow schools to determine the exact scope and timing of the classes themselves. Furthermore, the exact amount of instructional time varies across federal states and school types. Therefore, and acknowledging the interdisciplinary character of civic education, students were asked to answer the teaching quality instruments with regard to all classes in which they talk about political and civic affairs. A majority of participants ($n = 164$) were enrolled in an academic-track school (i.e. *Gymnasium*). The proportion of participants with an immigrant background was with 31.6% a little lower than in a comparable German study with 4th graders (Ohle-Peters et al. 2021, Mig. = 38.0%). In Germany, students have different teachers for different subjects. In this particular sample, students were part of a stable group of peers that remained intact across subjects.

This study was conducted in compliance with the German Research Foundation's guidelines for good scientific practice (*Deutsche Forschungsgemeinschaft, DFG*). In addition, the University Joint Ethics Committee examined and approved the research project with regard to protection of human dignity and the autonomy and self-determination of the people involved in the research. Prior to data collection, all participants and their parents were asked to give their informed consent. Each class was awarded 50 EUR for their participation.

Measures

Outcome: Willingness to Participate in Political and Civic Life

In order to assess their willingness to participate in political and civic life, students were given eleven items describing political and civic activities. Item selection was mainly driven by Ekman and Amnå's (2012) typology of political participation and prior empirical evidence (e.g., Albert et al., 2019; Gaiser et al., 2016). The introductory statement read as follows: "If you want to voice your opinion, how likely are you to take the following actions?" For each item, students chose a response on a four-point scale comprising 0 = "not at all likely", 1 = "rather not likely", 2 = "rather likely", and 3 = "very likely". Each item referred to a different political or civic activity. Thus, responses on each item were summed up to form an index ranging from 0 to 33, with higher scores indicating higher willingness to participate in more political and civic activities.

Teaching Quality Facets

In Germany, depending on grade level, federal state and school type, civic education classes differ with regard to their name, focus, degree of integration and position in the curriculum. In order to avoid misunderstandings, students were asked to answer the questions regarding teaching quality facets with respect to all situations in which they talk about civic or political affairs in class. Thus, students' perceptions of teaching quality cover but are not limited to civic education classes.

Open Classroom Climate. To capture students' perceptions of the frequency with which teachers created an open classroom climate, the German translation of the well-established scale from ICCS 2016 was used, e.g., "Teachers encourage students to make up their own minds" (Schulz, Carstens, et al., 2018). Specifically, students were provided with an introductory question asking "How often do you experience the following situations in class?" Students were then asked to rate each item on a four-point scale comprising 1 = "never", 2 = "rarely", 3 = "sometimes", and 4 = "often". In line with prior theoretical considerations and empirical evidence, the four items related to teachers' behaviors were used and the two items focusing on students' behaviors were disregarded. The scale's overall Cronbach's alpha reliability was .79 (.80 and .76 for 7th and 10th graders, respectively). For mean, standard deviation and correlations see Table 1.

Discussion of Current Political Events. Perceived frequency of discussions of current political events in class was measured with a single item asking "How often do you talk about

current political events in class?” (see e.g., Bittman & Russell, 2016). Again, responses were given on a 4-point scale from 1 = “never” to 4 = “often”.

Cognitive Activation. Perceived frequency of cognitive activation in class was measured with three items which were adopted from the COACTIV Study (Baumert et al., 2008, p. 129) and adjusted to the needs of this study. The items were preceded by the question “How often do you experience the following situations?” The three items were (1) “The teacher asks questions that cannot be easily answered, but which stimulate our thinking”, (2) “The teacher gives us tasks for which the primary goal is not to find the right solution but to find good arguments for our answer.”, and (3) “The teacher gives tasks for which we have to discuss with each other.” Each of the three items was rated on a four-point scale ranging from 1 = “never” to 4 = “often”. The scale’s overall alpha reliability was .79 (.81 and .76 for 7th and 10th graders, respectively).

Mediator A: Political Knowledge

To measure students’ political knowledge, an elaborate instrument was developed to match the skills of students in the 7th and 10th grade and the idiosyncrasies of the German political system (Alscher et al., 2022). Questions with different item difficulties were selected to represent the broad spectrum of political knowledge. A total of 99 items were used. The items were organized around the three content domains (1) political structures, (2) political processes, and (3) political principles dealing with different aspects of political and civic life, such as parliament, voting and human rights. Each item was a forced-choice question with three wrong and one correct answer option. Correct answers were coded as 1 and incorrect answers were coded as 0.

The 99 items included 36 grade-specific items for 7th and 10th grade as well as 27 anchor items answered by both 7th and 10th graders, resulting in a data structure known as a non-equivalent groups anchor test design (Paek & Cole, 2020). The difficulty of the items was approximated in advance of the surveys by a panel of experts consisting of professors of civic education instruction and study coordinators of large-scale assessment studies. Grade-specific items were assigned based on the difficulty levels determined by the experts. As students completed 63 items each, a maximum score of 63 points could be reached². To link the test scores

²For test development, we relied upon an advisory committee of experts, consisting of renowned German professors of civic education and the director of an international large-scale assessment of civic education and civic competence. The experts had prior experience with the development of knowledge tests. In a first phase, the experts were invited for a one-day workshop to talk about how the knowledge test framework could be conceptualized. In a second phase, the experts rated all preliminary items for plausibility, relevance, difficulty, and relationship to the knowledge test framework. Furthermore, they were given the opportunity to comment on the items in an open-ended way.

from 7th and 10th graders, a unidimensional multigroup 1-parameter model was calculated using the “TAM” package (Robitzsch et al., 2021) in R. Before calibration, rapid-guessing behavior was detected based on students’ response times. In accordance with a widely accepted approach, the bimodality of the response time distribution was used to visually identify the threshold at the local minimum between the two modes (Wise, 2017). For this test, this threshold was set slightly above the visual low point, at 6.5 seconds, leading to the deletion of 4.48% of answers. Including not-reached items, a total of 5.71% of responses were missing. During calibration, the estimated individual ability (θ) was calculated, ranging from -2.39 to 4.56, with higher scores representing more political knowledge. The WLE reliability was .88. For test security reasons, nine versions of each test, differing only in the order of the items, were administered.

Mediator B: Political Interest

Political interest was measured with the Short Scale Measuring Political Interest (SSPI; Otto and Bacherle, 2011). The scale includes five items, e.g., “I observe political events with great interest.” The introductory question read as follows: “To what extent do the following statements apply to you?” The students were asked to rate each item on a four-point scale ranging from 1 = “doesn’t apply at all” to 4 = “applies very well”. The scale’s overall alpha reliability was .89 (.87 and .91 for 7th and 10th graders, respectively).

Control Variables

Cultural capital was operationalized by the reported number of books in the student’s home (adapted from PISA 2015, see Mang et al., 2019). Responses ranged from 1 = “0 to 10 books” to 6 = “more than 500 books”. In addition to the numbers, each answer option was accompanied by a picture of a shelf displaying the relevant number of books. For immigrant background, students were asked which language they speak at home (adapted from IGLU 2016, see Hußmann et al. 2020). For those students who reported to speak at least sometimes a language other than German at home, an immigrant background was assumed.

Furthermore, 7th grade was coded as 0 and 10th grade as 1. In terms of school type, schools preparing for a vocational track were coded as 0 and schools preparing for an academic track were coded as 1. For school achievement, students’ self-reported math grade ranging from 1 = “inadequate” to 6 = “very good” was used.

Procedure

The study was administered in the first three class periods of a regular school day in the students' classrooms via computer-based assessment. After distributing the laptop computers to the students, trained research assistants explained the study procedure and the goals of the research project to the students. The study consisted of four parts: (1) a 60-minute political knowledge test, (2) a 20-minute cognitive ability test, (3) a 10-minute reading comprehension test, and (4) a 35-minute questionnaire. After each section, students were given short breaks. They could only resume their participation when the test coordinator provided them with the necessary keyboard shortcut to move on to the next screen on their laptop computers. After the data collection, students were given the opportunity to voice immediate feedback that will be used in future studies.

Data Analysis

All analyses were conducted using structural equation modelling and the “lavaan” package in R. The first structural equation model included the predictor variables (i.e., the civic education teaching quality facets) and the main outcome variable (i.e., students' willingness to participate in political and civic life). In the second model, the aforementioned control variables (i.e., cultural capital, immigrant background and school variables) were included as predictor variables to check the robustness of the results. In the third model, the mediating variables were added (i.e., political interest and political knowledge), while still enabling direct associations between the perceived civic education teaching quality facets and willingness to participate. In all models, open classroom climate, cognitive activation and political interest were operationalized as latent variables. All paths between the predictors, mediators and outcomes were specified. However, for the second model, only significant paths are displayed to declutter the figure. Intercorrelations were enabled between all exogenous variables and between the two mediating variables. For clarity, the intercorrelations of the control variables are not reported in the figures. A one-way random-effect analysis of the intra-class correlation revealed that only 4%, 6% and 5% of the variance in open classroom climate, cognitive activation and discussion of current political events, respectively, are explained by the class structure in the sample. Together with the small sample size, this led to the refusal of multilevel analysis.

We used the full information maximum likelihood (FIML) method to estimate the models with missing data. Moreover, the factor loadings of all latent variables' indicators were fixed to 1. Evaluation of the models' goodness of fit was based on the comparative fit index and

Tucker-Lewis index ($CFI \geq .95$ and $TLI \geq .90$) as well as the root mean square error of approximation ($RMSEA < 0.08$) and the standardized root mean square residual ($SRMR < 0.08$) (Hu & Bentler, 1999). Furthermore, the relative importance of the three civic education teaching quality facets was tested in both models by imposing equality constraints for the otherwise free parameters of the teaching quality facets. Afterwards, the model fits of the models with and without equality constraints were compared.

Results

Descriptive Statistics

The mean score for cultural capital (i.e., the number of books in the students' home) was a little higher than in comparable German studies (e.g., Heyder et al., 2021: $M = 2.88$) at $M = 3.55$. This could be explained by the marginal overrepresentation of students from higher-track schools in the sample. Average scores for the perceived civic education teaching quality facets were slightly above the midpoint of the response scale. This is in line with the German ICCS 2016 data (Abs & Hahn-Laudenberg, 2017) for perceived open classroom climate and with the 2010 NAEP data (Bittman & Russell, 2016) for perceived frequency of discussions of political events. The mean score for political interest, $M = 2.17$, is lower than the item-specific mean scores reported in the original pilot studies ($M = 2.50$ to 3.51 , Otto & Bacherle, 2011). However, only adults ($M_{age} = 38.6$ and 36.5) participated in these pilot studies, which presumably explains the higher average level of political interest. Comparable studies investigating German secondary education students' willingness to participate in political and civic life (Weißeno & Landwehr, 2018; Weißeno & Schmidt, 2019) have found that students demonstrate a moderate level of willingness to participate in political and civic life, which was in line with this study's results. The bivariate correlations in Table 1 show that the three facets of civic education teaching quality are statistically significantly, yet only moderately correlated with each other (.25 to .36).

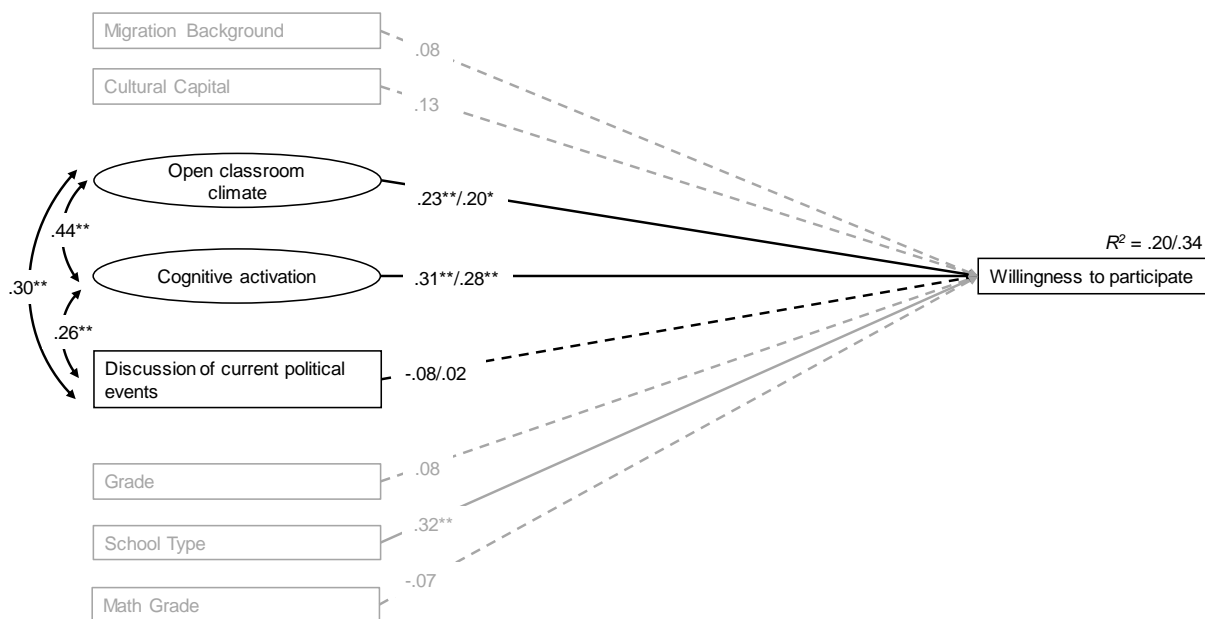
Associations Between Civic Education Teaching Quality and Students' Willingness to Participate in Political and Civic Life

The bivariate correlations show that perception of an open classroom climate and perceived cognitive activation are statistically significantly associated with students' willingness to participate in political and civic life (.31 and .35 respectively). Only the perceived frequency of discussions of current political events is not statistically significantly associated with students' willingness to participate. To examine how these associations might shift when considering all

three facets simultaneously, a partially latent structural equation model was fitted examining the direct associations between the perceived civic education teaching quality facets and students' willingness to participate. Looking at the standardized path coefficients (see Fig. 1), one sees that the perceived frequency of an open classroom climate and cognitive activation are indeed statistically significantly positively associated with students' willingness to participate in political and civic life ($\beta = .23$ and $.31$, $p = .007$ and $.000$ respectively). However, the model did not reveal a direct association between perceived frequency of discussions of current political events and students' willingness to participate ($\beta = -.08$, $p = .261$).

Figure 1

Direct associations between civic education teaching quality facets and willingness to participate.



Note. Values on the left are estimates without control variables, values on the right when control variables are included. Significant paths are solid; nonsignificant paths are dashed. CFI: .953/.955; TLI: .939/.923; RMSEA: .062/.054; SRMR: .052/.043.

* $p < .05$; ** $p < .01$

For the control variables, all associations with the outcome variable were in the expected direction, although only the association between school type and students' willingness to participate was statistically significant ($\beta = .32$, $p = .001$). The inclusion of the control variables led to only minor changes in the magnitude of the coefficients and little altered the significance

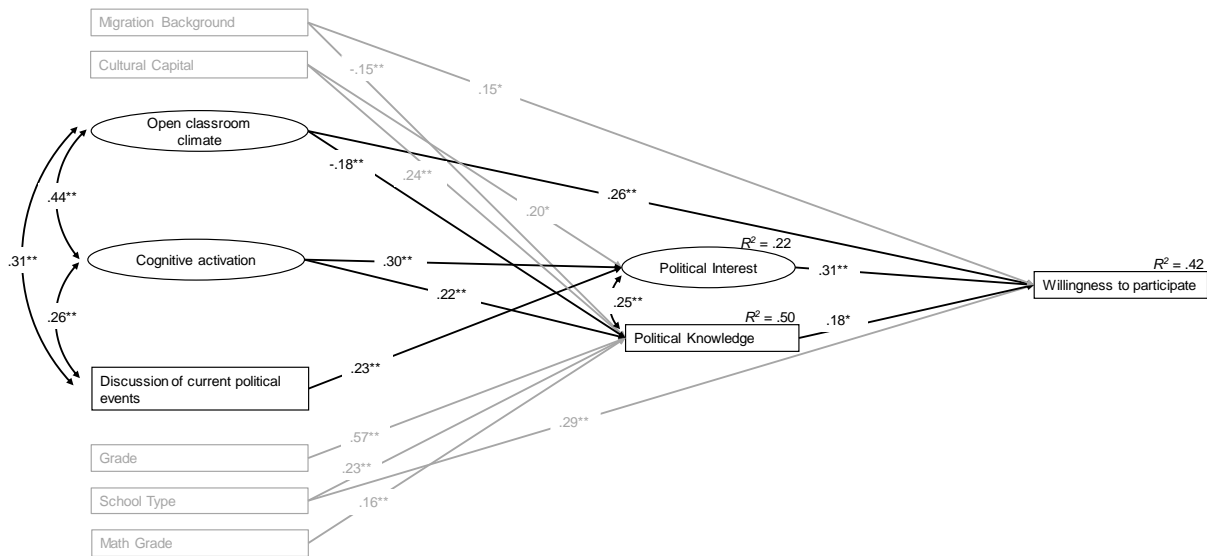
levels. Overall, the model explained 20% of the variance in the outcome variable, and 34% when including the control variables. For testing the relative importance of the teaching quality facets, the model that includes the control variables was used. The test showed that cognitive activation is a significantly stronger predictor for students' willingness to participate than the discussions of current political events. The chi-square test revealed that the model with equality constraints fits the data significantly worse than the model without equality constraints ($p = .020$). When comparing the importance of open classroom climate and the discussion of current political events, the chi-square test failed to meet the 5% significance threshold ($p = .080$). Moreover, cognitive activation was not a significantly stronger predictor than open classroom climate ($p = .686$).

Political Knowledge and Political Interest as Mediators between Teaching Quality and Students' Willingness to Participate in Political and Civic Life

To answer the second research question, a model was fitted examining the underlying mechanism of the associations between the civic education teaching quality facets and students' willingness to participate. For this purpose, political interest and political knowledge are included as mediating variables in the model (see Fig. 2). As expected, both political interest and political knowledge were statistically significantly positively associated with students' willingness to participate ($\beta = .31$ and $.18$, $p = .000$ and $.029$ respectively). Perceived frequency of an open classroom climate is still directly positively associated with willingness to participate ($\beta = .26$, $p = .001$). Rather counterintuitively, the perceived frequency of an open classroom climate is statistically significantly negatively associated with students' political knowledge ($\beta = -.18$, $p = .006$). Perceived frequency of cognitive activation is statistically significantly positively associated with both political interest ($\beta = .30$, $p = .001$) and political knowledge ($\beta = .22$, $p = .001$). The direct association between perceived frequency of cognitive activation and willingness to participate observed in the first model is now statistically significantly mediated by students' political interest ($\beta = .09$, $p = .006$). However, the mediation via political knowledge fails to meet the 5% significance threshold ($\beta = .04$, $p = .067$). The total indirect association is statistically significantly positive ($\beta = .13$, $p = .001$). Furthermore, perceived frequency of discussion of current political events is positively associated with political interest. The three education-related control variables grade, school type and math grade are associated with students' political knowledge but not with their political interest.

Figure 2

Associations between civic education teaching quality facets and willingness to participate mediated by political interest and political knowledge. Note: Only significant paths are displayed.



Note. CFI: .934; TLI: .910; RMSEA: .059; SRMR: .064.

* $p < .05$; ** $p < .01$

Overall, this model explains 42% of the variance in willingness to participate, 22% of the variance in political interest and 50% of the variance in political knowledge. The findings suggest that perceived frequency of an open classroom climate is directly associated with students' willingness to participate and not mediated by political knowledge or interest, while the association between perceived frequency of cognitive activation and students' willingness to participate is fully mediated by political interest and partially by political knowledge. Discussion of current political events is positively associated with political interest but neither directly nor indirectly associated with students' willingness to participate. To test for the relative importance, the total effects of teaching quality facets, thus both indirect effects through political interest and political knowledge as well as the direct effect, on students' willingness were compared. The chi-square test showed that cognitive activation is an overall stronger predictor than the discussion of current political events ($p = .018$). For the comparison of the open classroom climate and the discussion of current political events, the chi square test barely missed the 5% significance threshold ($p = .057$). Regarding the importance of cognitive activation and open classroom climate no significant differences were found ($p = .758$).

Sensitivity analyses and alternate model analyses

Additional analyses were carried out to confirm the robustness of the findings. Two separate analyses were run for participants from the 7th and the 10th grade. The overall pattern of findings remained intact. Numerically, school type plays a bigger role in the 10th grade while cultural capital and immigrant background are less important. The explained variance in students' willingness to participate was a little higher in the 10th than in the 7th grade. Regarding the civic education teaching quality facets, no systematic changes were observed. However, statistical power is limited due to the reduced sample sizes.

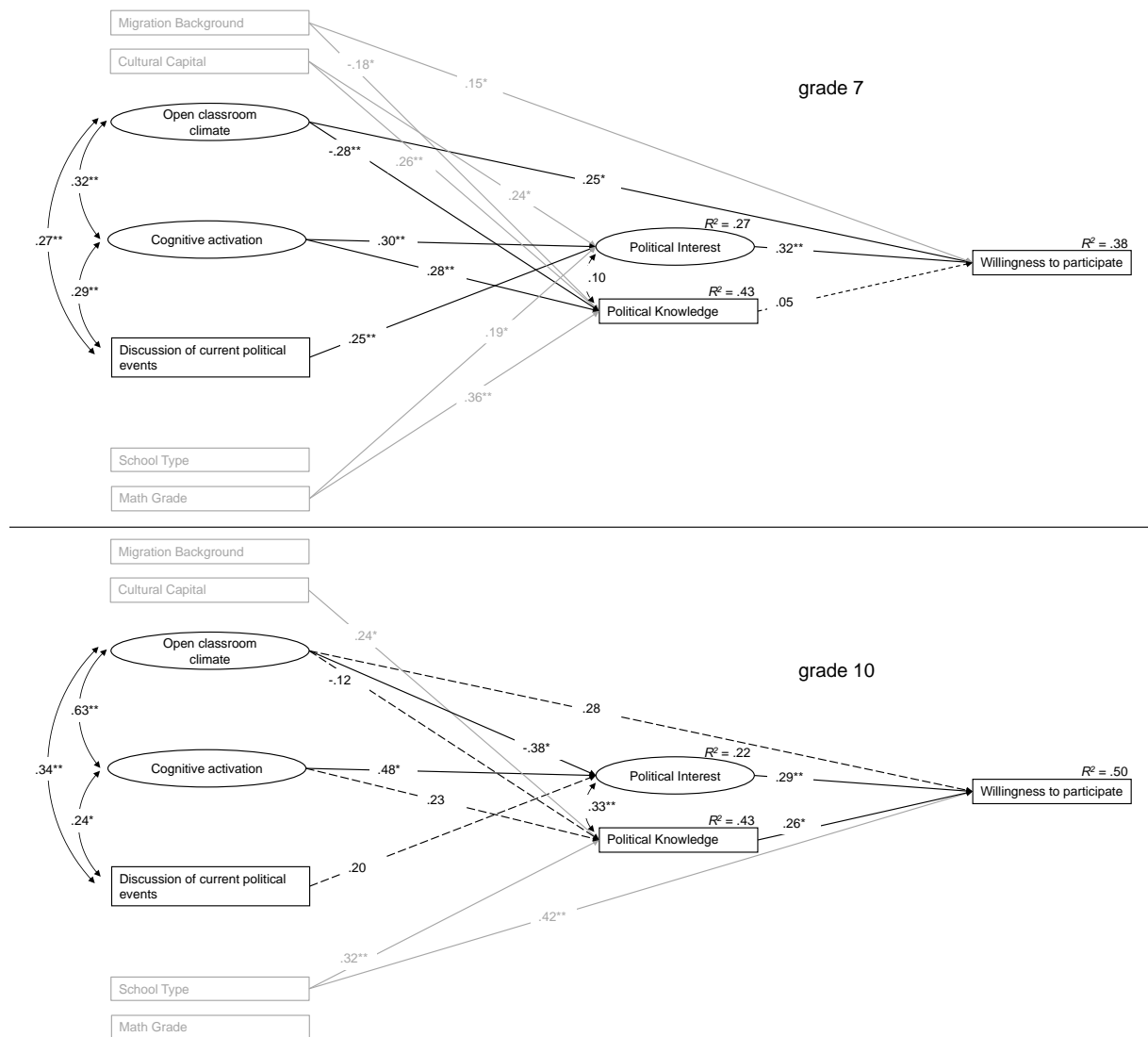
Discussion

Civic education is generally thought to be an important tool in educating youth to become responsible, informed and active citizens (e.g., Neundorf et al., 2016). In particular, civic education teaching quality facets and students' perceptions thereof play a crucial role in the achievement of desired civic outcomes among youth (Bayram-Özdemir et al., 2016; Manganelli et al., 2015). This is even more important in light of the limited time available for civic education in German schools' curriculum (Hedtke et al., 2021). However, previous research has usually focused on a single teaching quality facet, thus failing to test the relative importance of different teaching quality facets in civic education. Testing the relative importance matters because it enables us to better understand the contribution of different predictors to the variance in an outcome, unraveling the relative strength of these predictors. Moreover, whereas research on an open classroom climate and discussion of current political events have a long tradition in civic education research, cognitive activation has been largely overlooked. In addition, only little previous research has explored the mechanism underlying the relationship between civic education and willingness to participate by including mediating variables that are more immediate goals of civic education and function as prerequisites for political and civic participation.

The results show that a perceived open classroom climate and perceived cognitive activation are positively associated with students' willingness to participate. Students' perceptions regarding the frequency of discussions of current political events, however, are not associated with willingness to participate. The findings therefore only partially confirm the first hypothesis which assumes that all three facets have a positive relationship with willingness to participate. There is reason to believe in the robustness of these results, since the described associations remained after including control variables. The results regarding the relative importance of the facets further stress the importance of cognitive activation and open classroom climate.

Figure 3

Associations in 7th and 10th grade between civic education teaching quality facets and willingness to participate mediated by political interest and political knowledge.



Note. Only significant paths are displayed. G7: CFI: .925; TLI: .901; RMSEA: .060; SRMR: .076.; G10: CFI: .959; TLI: .943; RMSEA: .050; SRMR: .074.

* $p < .05$; ** $p < .01$

Political knowledge and political interest mediated the association between perceived cognitive activation and willingness to participate, suggesting that perceived cognitive activation plays an important role for the achievement of these more immediate goals of civic education. However, perceived open classroom climate still has a direct relationship with willingness to participate. Therefore, the findings again only partially confirm the second hypothesis. One reason for the continued direct relationship might be that the relationship is

mediated through another more immediate outcome of civic education, such as political efficacy (see Bayram-Özdemir et al., 2016; Marissen, 2020). Another explanation could of course be that an open classroom climate is indeed directly associated with willingness to participate. Future research should further investigate the mechanism underlying this relationship.

Furthermore, discussing current political events seems to be an important contributor to students' political interest, but not to their political knowledge. This finding is counterintuitive, at least given that previous findings suggest that the discussion of current political events is associated with increased political knowledge (Bittman's and Russel's 2016). However, unlike this study, Bittman and Russel (2016) did not control for an open classroom climate and cognitively activating teacher behavior. Data from the Civic Education Study in the United States in 1999 revealed that Traditional Teaching methods, including heavy usage of textbooks, memorization, worksheets, reports, extra materials and discussion of current political events has a statistically insignificant negative relationship with political knowledge (Martens & Gainous, 2013). However, the study also revealed that in combination with an open classroom climate, traditional teaching is the most effective overall pedagogical approach to civic education. One possible explanation for these inconsistent and sobering findings regarding the discussion of current political events in the classroom could be that teachers have varying time resources to prepare and lead discussions on current political events in a way that provides for meaningful learning opportunities (Kawashima-Ginsberg & Levine, 2014).

Surprisingly, the results show that when multiple teaching quality facets are considered a perceived open classroom climate is negatively associated with students' political knowledge. One possible explanation for this finding could be that more knowledgeable students have increased expectations regarding an open classroom climate and are therefore more often disappointed with actual classroom discussion conditions. A very open classroom might also come at the expense of structure and therefore hinder knowledge transfer. Another explanation might be that students who are very committed to academic achievement might feel a strong imperative to endorse their teacher's opinions and class material and hence, perceive the classroom to be less open. The bivariate correlation suggests no relationship between open classroom climate and political knowledge.

Other recent research also applies a more critical approach towards the assessment of open classroom climate. For instance, an open classroom climate alone does not ensure a nurturing social context in diverse classrooms (Munniksmä et al., 2021). In contradiction of previous findings (e.g., Campbell, 2008; Eckstein & Noack, 2015), an open classroom climate might not be effective in reducing social inequalities in political engagement (Deimel et al.,

2020; Hoskins et al., 2017). Moreover, a video study of nine 8th and 9th grade civic education classes in Germany, showed that open classroom discussions work differently depending on students' personality traits (Gronostay, 2019). For instance, agreeableness was negatively and extraversion positively associated with students' participation in class discussions. However, it should be mentioned that other recent studies did find positive, statistically significant associations between students' perceptions of an open classroom climate and political knowledge (e.g., Carrasco et al., 2020; Sampermans et al., 2021).

Moreover, the explained variance in students' willingness to participate in political and civic life nearly doubles in the second model compared to the first. This finding stresses the added value generated by introducing the more immediate goals of civic education as mediators. It should further be noted that the relationship between civic education teaching quality facets and political knowledge might be underestimated in this study. The knowledge test developed for this study was not tied to the curriculum and contains content that is learned outside of civic education classes. Hence, political learning outside of civic education classes and school might be more influential than it would be in a test that is tied to the curriculum. The results nonetheless suggest that political knowledge and political interest might be crucial intermediates in the relationship between civic education teaching quality and students' willingness to participate.

Despite its strengths, this study also comes with a few limitations. The data was collected in fall 2020, shortly before the second pandemic-driven lockdown in Germany, in which schools were closed entirely. The small sample size makes it impossible to account for the hierarchical structure of the data, since multilevel models did not converge. The low intra-class correlations indicate that students within a particular class evaluate teaching quality in their class differently. However, these findings are line with previous research (e.g., Praetorius, 2017) and might be explained by differences in perceptions or differences in actual treatment by the teacher. Future research should address the extent to which student perceptions, teacher statements, and objective measures (e.g., ratings by external observers) of instructional quality are interrelated and influence civic education outcomes. Furthermore, students were asked to answer the relevant questions with regard to all classes in which they talk about political and civic affairs, as not all students had dedicated civic education classes at the time data collection. This might underestimate the role of civic education instruction. It must be acknowledged that there may be discrepancies between students' perceptions of civic education teaching quality facets and the actual levels of these teaching quality facets. Ideally, future studies add to the literature by using independent measures of civic education teaching quality facets, such as

systematic class observations. Although previous studies have used a single-item solution to measure discussion of current political events (see Bittmann & Russel, 2016) and the corresponding findings yield plausible relationships, future studies should use a scale instead. Eventually, future research should aim for a better understanding of the relationship between an open classroom climate and students' political knowledge. There may also be additional facets of teaching quality in civic education that are not considered in this study, as well as other contextual influences, like participation in extracurricular civic activities, students' home environment, and students' peers. This study's strengths include its incorporation of multiple civic education teaching quality facets, making it possible to assess the relative importance of these facets. In addition, some of the instruments were newly developed in an elaborate, rigorous process and offer novel opportunities to measure perceived cognitive activation in civic education and political knowledge.

The results bear important implications for educational practice and policy. Due to their distinct mode of action, specific attention should be paid to different facets of teaching quality and their role in schools' fulfillment of their civic mission. Focusing on the creation of an open classroom climate alone might be insufficient or even detrimental for students' knowledge acquisition. Furthermore, vocational track students' disadvantageous position with regard to important civic outcomes warrants attention. There is thus a need for German educational policy to address the problem of systematic disparities in the formal qualification of civic education teachers across school tracks (NRW, 2020) which might be a source of this difference.

Conclusion

While in recent years, teaching quality in civic education has ranked highly on the scholarly agenda, the relative importance of different teaching quality facets remains understudied. To better understand teaching quality in civic education, the current study investigated three different teaching quality facets and their relationship with students' willingness to participate in political and civic life, their political knowledge and their interest. The results suggest that not all perceived teaching quality facets are equally important for students' willingness to participate. In particular, fostering and stimulating student engagement through cognitive activation seems to play a bigger role than previously assumed. In addition, the relationship between perceived cognitive activation and willingness to participate is mediated by political interest and political knowledge. Future research should continue to strive to better understand how teaching can improve civic education and contribute to schools' mission to produce responsible, informed and active citizens.

References

- Alscher, P., Ludewig, U., & McElvany, N. (2022). Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung. *Zeitschrift Für Erziehungswissenschaft*, 25(5), 1221–1241. <https://doi.org/10.1007/s11618-022-01085-0>
- Abs, H. J., & Hahn-Laudenberg, K. (2017). *Das politische Mindset von 14-Jährigen: Ergebnisse der International Civic and Citizenship Education Study 2016*. Waxmann.
- Achour, S., & Wagner, S. (2020). Ungleicher Zugang zur politischen Bildung: "Wer hat, dem wird gegeben". Ergebnisse einer bundesweiten Studie zur politischen Bildung und Demokratiebildung an Schulen. *Die Deutsche Schule*, 112(2), 143–158. <https://doi.org/10.25656/01:21702>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I., Timko, C., & White, J. (1982). Self-Monitoring and the Attitude-Behavior Relation. *Journal of Personality and Social Psychology*, 42, 426–435.
- Albert, M., Hurrelmann, K., Quenzel, G., & Kantar (Eds.). (2019). *18. Shell Jugendstudie: Jugend 2019 - eine Generation meldet sich zu Wort*. Beltz.
- Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.
- Barber, C., Torney-Purta, J., Wilkenfeld, B., & Ross, J. (2015). Immigrant and native-born adolescents' civic knowledge and attitudes in Sweden and the United States: Emergent citizenship within developmental niches. *Research in Comparative and International Education*, 10(1), 23–47. <https://doi.org/10.1177/1745499914567818>
- Barrett, M. (2018). *Kompetenzen für eine Demokratische Kultur: Gleichberechtigtes Zusammenleben in kulturell unterschiedlichen demokratischen Gesellschaften*. Sraßburg. Europarat.
- Baumert, J., Becker, M., Cortina, K. S., Köller, O., Kropf, M., & Maaz, K. (2016). Die Entwicklung des politischen Interesses und des Selbstkonzepts der politischen Kompetenz vom Jugend- bis in das Erwachsenenalter. In C. Grunert, A. Schippling, & N. Pfaff (Eds.), *Kritische Bildungsforschung: Standortbestimmungen und Gegenstandsfelder* (pp. 323–352). Barbara Budrich.
- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2008). *Professionswissen von*

- Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung: Nr. 83.* Max-Planck-Inst. für Bildungsforschung. <http://hdl.handle.net/hdl:11858/00-001M-0000-0023-998B-4>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Bayram-Özdemir, S., Stattin, H., & Özdemir, M. (2016). Youth's initiations of civic and political discussions in class: Do youth's perceptions of teachers' behaviors matter and why? *Journal of Youth and Adolescence*, 45(11), 2233–2245. <https://doi.org/10.1007/s10964-016-0525-z>
- Bittman, B., & Russell, W. (2016). Civic education in United States: A multiple regression of civic education scores from the National Assessment of Educational Progress. *Research in Social Sciences and Technology*, 1(2), 1–16. <https://doi.org/10.46303/ressat.01.02.1>
- Campaign for the Civic Mission of Schools. (2011). *Guardian of democracy: The civic mission of schools*. University of Pennsylvania, Annenberg Public Policy Center. https://media.carnegie.org/filer_public/ab/dd/abdda62e-6e84-47a4-a043-348d2f2085ae/ccny_grantee_2011_guardian.pdf
- Campbell, D. (2008). Voice in the Classroom: How an Open Classroom Climate Fosters Political Engagement Among Adolescents. *Political Behavior*, (30), 437–454.
- Carpini, M. X. D., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Carrasco, D., Banerjee, R., Treviño, E., & Villalobos, C. (2020). Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America. *Educational Psychology*, 40(2), 186–206. <https://doi.org/10.1080/01443410.2019.1699907>
- Castillo, J. C., Miranda, D., Bonhomme, M., Cox, C., & Bascopé, M. (2015). Mitigating the political participation gap from the school: the roles of civic knowledge and classroom climate. *Journal of Youth Studies*, 18(1), 16–35.
- Claes, E., & Hooghe, M. (Eds.) (2008). *Citizenship education and political interest: Political interest as an intermediary variable in explaining the effects of citizenship education*.

- Claes, E., & Hooghe, M. (2009). Citizenship Education and Political Interest. In *Civic Education and Youth Political Participation* (pp. 85–99). BRILL. https://doi.org/10.1163/9789460910258_006
- Claes, E., & Hooghe, M. (2017). The Effect of Political Science Education on Political Trust and Interest: Results from a 5-year Panel Study. *Journal of Political Science Education*, *13*(1), 33–45. <https://doi.org/10.1080/15512169.2016.1171153>
- Deimel, D., Hoskins, B., & Abs, H. J. (2020). How do schools affect inequalities in political participation: compensation of social disadvantage or provision of differential access? *Educational Psychology*, *40*(2), 146–166.
- Eckstein, K., Miklikowska, M., & Noack, P. (2021). School Matters: The Effects of School Experiences on Youth’s Attitudes toward Immigrants. *Journal of Youth and Adolescence*, *50*(11), 2208–2223. <https://doi.org/10.1007/s10964-021-01497-x>
- Eckstein, K., & Noack, P. (2015). Classroom climate effects on adolescents’ orientations toward political behaviors: a multilevel approach. In *Political Engagement of the Young in Europe* (pp. 181–197). Routledge.
- Ekman, J., & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, *22*(3), 283–300. <https://doi.org/10.2478/s13374-012-0024-1>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Förtsch, C., Werner, S., Kotzebue, L. von, & Neuhaus, B. J. (2016). Effects of biology teachers’ professional knowledge and cognitive activation on students’ achievement. *International Journal of Science Education*, *38*(17), 2642–2666. <https://doi.org/10.1080/09500693.2016.1257170>
- Gaiser, W., Hanke, S., & Ott, K. (2016). *Jung-politisch-aktiv?! Politische Einstellungen und politisches Engagement junger Menschen.: Ergebnisse der FES-Jugendstudie 2015*. Dietz Verlag.
- Galston, W. (2001). Political knowledge, political engagement, and civic education. *Annual Review of Political Science*, *4*, 217–234.
- Gil de Zúñiga, H., & Diehl, T. (2019). News finds me perception and democracy: Effects on political knowledge, political interest, and voting. *New Media & Society*, *21*(6), 1253–1271.

- Godfrey, E. B., & Grayman, J. K. (2014). Teaching citizens: The role of open classroom climate in fostering critical consciousness among youth. *Journal of Youth and Adolescence*, *43*(11), 1801–1817.
- Goodin, R. E., & Spiekermann, K. (2018). *An epistemic theory of democracy*. Oxford University Press.
- Grobshäuser, N., & Weißeno, G. (2020). Does political participation in adolescence promote knowledge acquisition and active citizenship? *Education, Citizenship and Social Justice*, *16*(2), 150–164. <https://doi.org/10.1177/1746197919900153>
- Gronostay, D. (2019). To argue or not to argue? The role of personality traits, argumentativeness, epistemological beliefs and assigned positions for students' participation in controversial political classroom discussions. *Unterrichtswissenschaft*, *47*(1), 117–135. <https://doi.org/10.1007/s42010-018-00033-4>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Hedtke, R., Gökbudak, M., & Hagedorn, U. (2021). 4. Ranking Politische Bildung Politische Bildung in der Sekundarstufe I und in der Berufsschule im Bundesländervergleich 2020. Didaktik der Sozialwissenschaften: Working Papers 12.
- Hess, D. E., & McAvoy, P. (2015). *The political classroom: Evidence and ethics in democratic education*. *Critical social thought*. Routledge. <https://doi.org/10.4324/9781315738871>
- Heyder, A., Weidinger, A. F., & Steinmayr, R. (2021). Only a Burden for Females in Math? Gender and Domain Differences in the Relation Between Adolescents' Fixed Mindsets and Motivation. *Journal of Youth and Adolescence*, *50*(1), 177–188. <https://doi.org/10.1007/s10964-020-01345-4>
- Hoskins, B., Janmaat, J. G., & Melis, G. (2017). Tackling inequalities in political socialisation: A systematic analysis of access to and mitigation effects of learning citizenship at school. *Social Science Research*, *68*, 88–101. <https://doi.org/10.1016/j.ssresearch.2017.09.001>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.
- Hußmann, A., Wendt, H., Bos, W., & Rieser, S. (Hrsg.) (2020). *IGLU 2016. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen* (1st, New ed.). Waxmann.
- Inglehart, R., & Welzel, C. (2003). Political culture and democracy: Analyzing cross-level linkages. *Comparative Politics*, 61–79.

- Isac, M. M., Maslowski, R., & van der Werf, G. (2011). Effective civic education: an educational effectiveness model for explaining students' civic knowledge. *School Effectiveness and School Improvement*, 22(3), 313–333. <https://doi.org/10.1080/09243453.2011.571542>
- Janmaat, J. G., Mostafa, T., & Hoskins, B. (2014). Widening the participation gap: The effect of educational track on reported voting in England. *Journal of Adolescence*, 37(4), 473–482. <https://doi.org/10.1016/j.adolescence.2014.03.011>
- Kahne, J. (2005). *Educating for Democracy: The California Campaign for the Civic Mission of Schools*. The California Survey of Civic Education.
- Kahne, J., Crow, D., & Lee, N.-J. (2013). Different Pedagogy, Different Politics: High School Learning Opportunities and Youth Political Engagement. *Political Psychology*, 34(3), 419–441. <https://doi.org/10.1111/j.1467-9221.2012.00936.x>
- Kawashima-Ginsberg, K., & Levine, P. (2014). Diversity in classrooms: The relationship between deliberative and associative opportunities in school and later electoral engagement. *Analyses of Social Issues and Public Policy*, 14(1), 394–414.
- Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409.
- Kultusministerkonferenz. (2018). *Demokratiebildung*. <https://www.kmk.org/themen/allgemein-bildende-schulen/weitere-unterrichtsinhalte/demokratiebildung.html>
- Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A., & Reiss, K. (2019). *PISA 2015 Skalendhandbuch: Dokumentation der Erhebungsinstrumente*. Waxmann. <https://elibrary.utb.de/doi/book/10.31244/9783830990321>
- Manganelli, S., Alivernini, F., Lucidi, F., & Di Leo, I. (2012). Expected Political Participation in Italy: a Study based on Italian ICCS Data. *Procedia - Social and Behavioral Sciences*, 46, 1476–1481. <https://doi.org/10.1016/j.sbspro.2012.05.324>
- Manganelli, S., Lucidi, F., & Alivernini, F. (2014). Adolescents' expected civic participation: The role of civic knowledge and efficacy beliefs. *Journal of Adolescence*, 37(5), 632–641. <https://doi.org/10.1016/j.adolescence.2014.05.001>
- Manganelli, S., Lucidi, F., & Alivernini, F. (2015). Italian adolescents' civic engagement and open classroom climate: The mediating role of self-efficacy. *Journal of Applied Developmental Psychology*, 41, 8–18. <https://doi.org/10.1016/j.appdev.2015.07.001>
- Martens, A. M., & Gainous, J. (2013). Civic education and democratic capacity: How do teachers teach and what works? *Social Science Quarterly*, 94(4), 956–976.

- Maurissen, L. (2020). Political efficacy and interest as mediators of expected political participation among Belgian adolescents. *Applied Developmental Science, 24*(4), 339–353. <https://doi.org/10.1080/10888691.2018.1507744>
- McCafferty-Wright, J., & Knowles, R. (2016). Unlocking the Civic Potential of Current Events with an Open Classroom Climate. *Social Studies Research & Practice (Board of Trustees of the University of Alabama), 11*(3).
- McWhirter, E. H., & McWhirter, B. T. (2016). Critical Consciousness and Vocational Development Among Latina/o High School Youth: Initial Development and Testing of a Measure. *Journal of Career Assessment, 24*(3), 543–558.
- Milner, H. (2002). *Civic Literacy: How Informed Citizens Make Democracy Work*. University Press of New England. <https://books.google.de/books?id=zy2GAAAAMAAJ>
- Munniksmma, A., Ziemes, J., & Jugert, P. (2021). Ethnic Diversity and Students' Social Adjustment in Dutch Classrooms. *Journal of Youth and Adolescence*. Advance online publication. <https://doi.org/10.1007/s10964-021-01507-y>
- Neundorf, A., Niemi, R. G., & Smets, K. (2016). The compensation effect of civic education on political engagement: How civics classes make up for missing parental socialization. *Political Behavior, 38*(4), 921–949.
- Niemi, R. G., & Junn, J. (1998). *Civic education: What makes students learn*. Yale University Press. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=52882>
- Ohle-Peters, A., Iglar, J., Schlitter, T., Teerling, A., Köller, O., & McElvany, N. (2021). Unterrichtsqualität und intrinsische Lesemotivation im Kontext der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ (BiSS). *Zeitschrift für Erziehungswissenschaft, 24*(4), 861–882.
- Otto, L., & Bacherle, P. (2011). Politisches Interesse Kurzsкала (PIKS): Entwicklung und Validierung. *Politische Psychologie, 1*(1), 19–35.
- Paek, I., & Cole, K. (2020). *Using R for item response theory model applications*. Routledge. <https://doi.org/10.4324/9781351008167>
- Prior, M., & Bougher, L. D. (2018). “Like they’ve never, ever seen in this country”? Political interest and voter engagement in 2016. *Public Opinion Quarterly, 82*(1), 822–842.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.

- Quintelier, E., & Hooghe, M. (2012). Political attitudes and political participation: A panel study on socialization and self-selection effects among late adolescents. *International Political Science Review*, 33(1), 63–81. <https://doi.org/10.1177/0192512111412632>
- Quintelier, E., & Hooghe, M. (2013). The relationship between political participation intentions of adolescents and a participatory democratic climate at school in 35 countries. *Oxford Review of Education*, 39(5), 567–589. <https://doi.org/10.1080/03054985.2013.830097>
- Reichert, F. (2016). How Internal Political Efficacy Translates Political Knowledge Into Political Participation: Evidence From Germany. *Europe's Journal of Psychology*, 12(2), 221–241. <https://doi.org/10.5964/ejop.v12i2.1095>
- Reichert, F., & Torney-Purta, J. (2019). A cross-national comparison of teachers' beliefs about the aims of civic education in 12 countries: A person-centered analysis. *Teaching and Teacher Education*, 77, 112–125. <https://doi.org/10.1016/j.tate.2018.09.005>
- Robitzsch, A., Kiefer, Thomas, Wu, & Margaret. (2021). *TAM: Test analysis module*. <https://cran.r-project.org/web/packages/TAM/index.html>
- Roccatto, M., & Zogmaister, C. (2010). Predicting the Vote through Implicit and Explicit Attitudes: A Field Research. *Political Psychology*, 31(2), 249–274. <http://www.jstor.org/stable/20721288>
- Sampermans, D., Claes, E., & Janmaat, J. G. (2020). Back on track? How civic learning opportunities widen the political knowledge gap in a tracked education system. *School Effectiveness and School Improvement*, 32(2), 241–259. <https://doi.org/10.1080/09243453.2020.1830125>
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2009). *ICCS 2009 International Report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries*. Springer. http://www.erc.ie/documents/iccs2009_international_report.pdf
- Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G., & Friedman, T. (2018). *Becoming Citizens in a Changing World: IEA International Civic and Citizenship Education Study 2016: International Report*. Springer. <https://doi.org/10.1007/978-3-319-73963-2>
- Schulz, W., Carstens, R., Losito, B., & Fraillon, J. (2018). *ICCS 2016 Technical Report: IEA International Civic and Citizenship Education Study 2016*. International Association for the Evaluation of Educational Achievement.
- Stang, J., & McElvany, N. (2020). Unterschiede in der Wahrnehmung der Qualität des Deutschunterrichts zwischen Grundschülerinnen und Grundschülern. *Zeitschrift Für Pädagogische Psychologie*, 35, 261–273. <https://doi.org/10.1024/1010-0652/a000275>

- Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at age fourteen*. The International Association for the Evaluation of Educational Achievement. <http://www.wam.umd.edu/%7Eiea/>
- Wallrich, L., West, K., & Rutland, A. (2021). Civic understanding mediates of the effect of educational tracks on voting intentions in German schools. *Citizenship, Social and Economics Education*, 20(2), 84–102. <https://doi.org/10.1177/20471734211016501>
- Weißeno, G., & Landwehr, B. (2015). Knowledge about the European Union in Political Education: What are the Effects of Motivational Predispositions and Cognitive Activation? *McGill Journal of Education / Revue Des Sciences De L'éducation De McGill*, 50(2-3), 413–432.
- Weißeno, G., & Landwehr, B. (2018). Politische Partizipation, Selbstkonzept und Fachwissen: Ergebnisse einer Studie. In B. Ziegler & M. Waldis (Eds.), *Politische Bildung in der Demokratie: Interdisziplinäre Perspektiven* (pp. 175–190). Springer. https://doi.org/10.1007/978-3-658-18933-4_12
- Weißeno, G., & Schmidt, A. (2019). Fachinteresse, Selbstwirksamkeit, Partizipationsbereitschaft und politisches Wissen von Oberstufenschüler/-innen. In G. Weißeno (Ed.), *Politik lernen: Studien und theoretische Ansätze* (pp. 195–211). Springer. https://doi.org/10.1007/978-3-658-27896-0_12
- Wise, S. (2017). Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications: Its Identification, Interpretation, and Implications. *Educational Measurement: Issues and Practice*, 36, 52–61.
- Yang, J., & Hoskins, B. (2020). Does university have an effect on young people's active citizenship in England? *Higher Education*, 80(5), 839–856. <https://doi.org/10.1007/s10734-020-00518-1>

3 Weiterführende Analysen

3.1 When Will They Know What They Don't Know? – Political knowledge and the Infamous Dunning-Kruger Effect

Alscher, P., Ludwig, U., Kleinkorres, R. & McElvany, N. (2023). When Will They Know What They Don't Know? – Political knowledge and the Infamous Dunning-Kruger Effect [Manuskript zur Publikation eingereicht]. TU Dortmund.

Dieser Artikel stellt keine Kopie der Veröffentlichung dar und entspricht möglicherweise nicht exakt der endgültigen und maßgeblichen Version des in der Zeitschrift veröffentlichten Artikels.

Abstract

The Dunning-Kruger effect describes that low performers tend to overestimate and high performers tend to underestimate their ability and that on average low performers provide less accurate estimates of their ability than high performers. The classical interpretation of the Dunning-Kruger effect postulates that low performers suffer from a deficiency in metacognitive insight, while critiques argue that the Dunning-Kruger effect is merely a statistical artifact. Based on data from $N = 1,047$ students in Grade 7 ($n = 613$) and Grade 10 ($n = 434$), we examined whether the Dunning-Kruger effect exists with regard to high school students' political knowledge, whether the effect is different in Grade 7 and Grade 10, and how teachers' behavior (i.e., cognitively activating and motivating teaching) affects students' judgement accuracy and judgement direction. The results show that the Dunning-Kruger effect exists in both grade levels. Furthermore, visual and statistical examination of the data suggest that the pattern of data is very similar in both grade levels. Finally, we further find that perceived motivational quality is associated with judgement accuracy, but also with overconfidence. Perceived cognitive activation is negatively associated with judgement accuracy. Through the usage of two-cohort data, this study applies a novel approach and bears important implications for research on the Dunning-Kruger effect. Finding no differences between the two cohorts at different levels of development lends support to the notion that the Dunning-Kruger effect might be a statistical artifact. Furthermore, the results regarding the teachers' behavior provide important insights for educational practice in civic education.

Keywords: Adolescence · Dunning-Kruger effect · Metacognitive insight · Self-assessment · Political Knowledge

Introduction

How accurately individuals evaluate their own ability and performance can be crucial in several contexts. For instance, overconfidence can impede effective learning and produce underachievement (Dunlosky & Rawson, 2012). Furthermore, overconfidence has repeatedly been shown to be associated with radical political beliefs (e.g., van Prooijen, 2021). Of course, overconfidence can have positive consequences, too. For example, overconfidence is associated with increased social success over time in at least some domains (Murphy et al., 2018). The status-enhancement theory stipulates that confident individuals attain higher status and are viewed positively by their peers regardless of whether their confidence is justified by actual ability (Kennedy et al., 2013). On the contrary, underconfidence is theorized to limit an individual's motivation (Bandura, 1977). Indeed, empirical evidence suggests that underconfidence in one's own ability is associated with psychological maladjustment, low motivation and low educational achievement (Bouffard et al., 2011; Sheldrake et al., 2022).

Regarding self-assessed ability, empirical evidence suggests that low performers' estimates are less accurate than high performers' estimates and that low performers have a tendency for overestimation while high performers have a tendency for underestimation (Kruger & Dunning, 1999). Named after the authors, who conducted multiple experiments on the relationship between self-assessed and objective ability, this pattern has become known as the Dunning-Kruger effect. Kruger and Dunning (1999) argued that low performers suffer from a 'dual curse' (see Schlösser et al., 2013). The dual curse postulates that low performers tend to be not only incompetent with respect to a specific domain, but also lack the metacognitive insight to acknowledge their own incompetence. In the educational context, a lack of metacognitive insight can result in learning deficits going unrecognized and compromising learning and educational success (Dunlosky et al., 2021). However, critics argue that the Dunning-Kruger effect is merely a statistical artifact primarily caused by the better-than-average effect and regression-to-the-mean (e.g., Gignac & Zajenkowski, 2020).

Using the example of performance on a political knowledge test, we investigated the Dunning-Kruger effect and added empirical evidence to the question of whether the Dunning-Kruger effect is a statistical artifact. Exploring the Dunning-Kruger effect within the context of political knowledge might be particularly interesting, as individuals are constantly exposed to political content even when they do not seek it out. Furthermore, we examined the Dunning-Kruger effect from an educational perspective by examining the relationship between teaching quality and judgement accuracy as well as judgement direction.

For this purpose, we used data from high school students in Germany. The dynamics and plasticity of metacognitive abilities in adolescence make high school an interesting case study. Specifically, we investigated students' performance in a political knowledge test and their self-assessed political knowledge in Grade 7 and Grade 10. Thereby, the present study provides a novel approach to research on the Dunning-Kruger effect. The study of two cohorts that are at least theoretically different in terms of average metacognitive abilities (in this case students from two different grade levels) allows for an original assessment of the Dunning-Kruger effect and adds empirical evidence to the question of whether it is a statistical artifact or whether metacognitive insight is responsible for it. In addition, the educational context is considered by addressing teaching quality as a key determinant of the learning environment and its relationship to students' judgements.

The Dunning-Kruger effect

In the past decades, empirical research in numerous fields has reported evidence that self-assessed ability (e.g., intelligence, language competence, math competence) and objectively measured ability are only moderately associated. For instance, a meta-analysis including 55 studies with a total of 267 correlation coefficients found a low to moderate correlation between self-assessed ability and different measures of performance (i.e., objective tests, class grades and supervisor ratings), with high variability ($r = .29$, $SD = .25$; Mabe & West, 1982). A similar meta-analysis 30 years later with 154 effect sizes estimated the relationship between self-estimated and psychometrically measured intelligence at $r = .33$ (Freund & Kasten, 2012). In addition, in a metasynthesis, the average correlation between self-assessed ability and objectively measured ability across 22 meta-analyses in various domains was $r = .29$ (Zell & Krizan, 2014). Depending on the domain, the magnitude of correlation coefficients varies from $r \approx .20$ for verbal tasks to $r \approx .40$ for math (Syzmanowicz & Furnham, 2011). Moreover, correlations become larger (i.e., self-assessments become more accurate) in cases of test expectation, test experience, relative self-assessments, and anonymous self-assessments (Ackerman & Wolman, 2007; Mabe & West, 1982).

The Dunning-Kruger effect describes a pattern of systematic difference in the relationship between self-assessed and objectively measured performance. In their original study, Kruger and Dunning (1999) used data from Cornell University undergraduates and examined students' performances in the fields of grammar, logical reasoning and humor. They found the same pattern in all three domains: low performers tended to overestimate their performance, high performers tended to underestimate their performance, and high performers' estimates

were more accurate than low performers' estimates. The effect was subsequently found in various other fields and domains. For instance, the results of a grammar study with $N = 3,515$ participants and a logical reasoning study with $N = 3,543$ participants recruited from Amazon's Mechanical Turk each affirmed the general pattern of the Dunning-Kruger effect and the notion that low performers are less able to provide accurate estimates of their ability than high performers (Jansen et al., 2021). An investigation of $N = 150$ Türkiye-based tourism and hospitality employees' self-reports of cultural awareness and their performance in the Cross-Cultural Awareness Test showed that low performers were especially prone to overestimations, providing higher self-reports of cultural awareness than individuals with middling levels of performance (Koc, 2021). An examination of $N = 94$ Canadian high school volleyball coaches' self-assessed coaching efficacy and objectively measured coaching ability revealed that low performers had higher efficacy than ability and high performers had lower efficacy than ability (Sullivan et al., 2018). However, while viewed as evidence supporting the 'generalizability of the Dunning-Kruger effect', the study does not provide evidence that high performers' estimates are more accurate than low performers' estimates.

The sheer continuity with which the effect has been found has spurred some scholars to try to find the 'special' domain in which it does not arise (Zhou & Jenkins, 2020). Even scholars who criticize Kruger's and Dunning's interpretation of the effect (i.e., that low performers lack the metacognitive insight to acknowledge their incompetence) and view the effect as a statistical artifact (i.e., caused by the better-than-average effect and regression-to-the-mean) find the typical pattern of the Dunning-Kruger effect when replicating the study (e.g., Gignac & Zajenkowski, 2020).

The role of metacognitive insight in the Dunning-Kruger effect

The Dunning-Kruger effect leads scholars to believe that low performers suffer from a dual curse of (1) not being competent at a task and (2) not being capable of acknowledging their incompetence at the task (e.g., Jansen et al., 2021; Rahmani, 2020; for a contrary view arguing that the effect is a statistical artifact, see for instance Gignac & Zajenkowski, 2023; McIntosh et al., 2022). In their original work, Kruger and Dunning attributed the second curse to a deficit in metacognition among low performers. In broad terms, metacognition can be differentiated into two main components: metacognitive knowledge (i.e., 'thinking about thinking') and metacognitive experiences (i.e., 'managing thinking', 'monitoring performance') (Flavell, 1979). Accordingly, metacognition focusses on the process of problem solving, whereas cognition centers on actually solving the problem (Marchant, 1989).

In particular, Kruger and Dunning (1999) claimed that poor performers lack metacognitive insight, which is a subcomponent of metacognitive experiences. The argument goes that for any domain, metacognitive insight requires the very same skill as task performance. Hence, it takes competence to see lack of competence (Dunning, 2011).

Albeit receiving generally less attention, the pattern of high performers tending to underestimate their competence could be explained by ‘undue modesty’ (e.g., Dunning et al., 2003). Another explanation holds that estimation errors among poor performers involve misvaluations of their own competence, while estimation errors among high performers are based on misvaluations of the competence of others (Schlösser et al., 2013). Indeed, counterfactual regression analyses provided evidence that low performers’ estimation errors were primarily due to errors in estimating their own performance (Ehrlinger et al., 2008). On the contrary, high performers’ estimation errors were due to both errors in estimating their own performance and errors in estimating others’ performances, indicating that they partially fall victim to a false consensus effect (Ehrlinger et al., 2008).

In the original paper by Kruger and Dunning (1999), one study included a measure of metacognition. After participants had answered a 10-item logical reasoning task and given global ratings of how they thought they had performed, they were asked to indicate for each item individually whether they thought they answered it correctly. Indeed, this measure correlated strongly with task performance, indicating that low performers seem to be at least partially aware of their knowledge gaps. However, critics argue that this relationship is influenced by an optimism bias (see Krueger & Mueller, 2002).

In two studies investigating the relationship between metacognitive ability and accuracy of ability self-estimates, $N = 91$ and $N = 113$ students from a cognitive psychology course at Texas A&M University were asked to estimate their exam grade and their confidence in their estimate (Miller & Geraci, 2011). Again, low-performing students were more likely to overestimate their exam score, but they were also less confident about their grade estimations. Furthermore, the study showed that the likelihood of overestimation decreased from predictions to postdictions. Finally, a study based on responses of $N = 101$ students and alumni of the University of Edinburgh showed that metacognitive skills, measured as metacognitive sensitivity and calibration, can contribute to the Dunning-Kruger effect but are neither necessary nor sufficient for it (McIntosh et al., 2019).

Alternative explanations for the Dunning-Kruger effect

Not long after the original paper's publication, the classic metacognitive interpretation of the Dunning-Kruger effect was challenged by alternative explanations. Critics argued that the effect is merely a statistical artifact and not due to a systematic lack of metacognitive insight among low performers (e.g., Krueger & Mueller, 2002; for a reply, see Kruger & Dunning, 2002). These critiques state that the reason for overestimation among low performers is basically twofold: regression-to-the-mean and the better-than-average effect.

Regression-to-the-mean describes the phenomenon that an individual with an extreme value in variable A is likely to have a less extreme value in variable B (Gignac & Zajenkowski, 2020). The argument goes that for any measure involving some degree of chance, such as complex phenomena with multiple causes, extreme values tend to be associated with less extreme values. Hence, whenever the correlation between two scores is imperfect (i.e., $|r| < 1.0$), there will be some regression-to-the-mean. Kruger and Dunning (1999) believed that despite the inevitability of regression-to-the-mean, the reasons for overestimation among low performers primarily lay in psychological rather than statistical mechanisms. They argued that if regression-to-the-mean was to blame for their results, the magnitude of miscalibration among low performers should not have exceeded that of high performers. This assumption was later at least partially invalidated by the better-than-average effect (e.g., Gignac & Zajenkowski, 2020).

The better-than-average effect describes individuals' tendency to rate their own performance or ability as above average (Zell et al., 2020). Unless a distribution is negatively skewed, only half the distribution's population can be above average on a certain characteristic. However, when it comes to abilities, performances or traits, a majority of people believe they are above average. A meta-analysis of 291 samples supports the robustness of the better-than-average effect ($d_z = 0.78$, 95% CI [0.71, 0.84]) across different domains (Zell et al., 2020). Together, regression-to-the-mean and the better-than-average effect provide a sound explanation for why low performers tend to overestimate and high performers tend to underestimate their ability, with the overestimation exceeding the underestimation in magnitude.

Another, less prominent criticism centers on the generalizability of the original results. Krajč and Ortmann (2008) argue that Kruger and Dunning used convenience samples – usually comprising undergraduate (psychology) students from Cornell University – rather than representative samples. Based on Cornell University's selectivity, the samples used by Kruger and Dunning were not distributed uniformly with regard to performance, but were skewed toward the bottom, forming a backward-*J* distribution. Accordingly, Krajč and Ortmann (2008) argued that the structure of the data would make it particularly difficult for relatively low performers

to extract the signal of ability level when placing themselves relative to their peers. However, the Dunning-Kruger effect has been replicated in less competitive contexts, with more representative samples in various settings, and with different data structures (e.g., Schlösser et al., 2013).

The claim that incentivizing accurate estimations through financial and social rewards could lead to more accurate estimations among all performers was preemptively refuted with a series of studies indicating that neither financial nor social incentives influenced overestimation among low performers (Ehrlinger et al., 2008).

Metacognitive ability in education

Adolescence is a life period associated with dynamic shifts in cognitive (Galambos et al., 2005), social cognitive (Choudhury et al., 2006), and metacognitive ability (Weil et al., 2013). Research has demonstrated that metacognitive ability is a key variable for students' development and academic achievement (e.g., Ohtani & Hisasaka, 2018; Perry et al., 2019). It has been recognized as an important factor driving students' learning (Baars et al., 2018; Davis & Chan, 2023; Metcalfe, 2009).

From a developmental point of view, early adolescence is a particularly important life phase regarding the attainment of metacognitive ability. For instance, a longitudinal study with $N = 32$ secondary school students showed qualitative and quantitative growth in metacognitive ability from age 12 to age 14 (Van der Stel & Veenman, 2010). In addition, the study showed that metacognitive ability contributes to students' learning performance somewhat independently of intellectual ability. A follow-up study showed that the development of metacognitive ability significantly slows down at age 15 (Van der Stel & Veenman, 2014). A study with $N = 133$ third-graders and $N = 73$ fifth-graders indicated that the fifth-graders were better able to improve their test performance by withdrawing potentially wrong answers when given the chance, hence demonstrating better metacognitive ability (Roebbers et al., 2009). In addition, a longitudinal study (16 months, two measurement points) with $N = 853$ fifth-graders (at the first measurement point) in Germany, which investigated metacognitive knowledge in the domains of German and English, found statistically significant increases in students' metacognitive knowledge in both domains and across school tracks (Artelt et al., 2012). Under certain circumstances, similar results can be found for older cohorts. A one-year panel study with $N = 66$ undergraduates between the age of 20 and 25 in Hong Kong demonstrated that metacognitive ability can increase even at a higher age (Downing et al., 2009). Students who completed a course of study and assessment rooted exclusively in a problem-based learning

approach experienced statistically significant increases in their metacognitive ability, whereas students in a program with a distinctly non-problem-based learning approach experienced hardly any changes in their metacognitive ability. Finally, a study with cross-sectional data from the United Kingdom examined objective performance on a visual task and metacognitive ability among $N = 56$ participants between 11 and 41 years of age (Weil et al., 2013). Metacognitive ability improved throughout adolescence, plateaued going into adulthood and slowly fell afterwards. Generally speaking, study results suggest that metacognitive ability increases during early and late childhood (e.g., Artelt et al., Flavell et al., 1998), that the increase slows down during adolescence (e.g., dos Santos Kawata; Weil et al., 2012) and that metacognitive ability reaches a plateau when entering adulthood (e.g., Flavell et al., 1998; Weil et al., 2013).

A driving force for the attainment of metacognitive ability is school education and teachers' behavior. Although difficult to teach and slow in its development, metacognitive ability and the application of metacognitive strategies can be taught in schools and developed with practice (Dignath & Büttner, 2008). A one-year case study of a fourth-grade class demonstrated that teachers can support the acquisition of metacognitive ability through the use of structured thinking activities (Branigan & Donaldson, 2020). Of particular importance in this context is that teachers create a classroom that values students' perspectives, make metacognitive thinking processes visible by acting as 'metacognitive role models', and provide enough time and opportunities for students to think about and scrutinize their own thinking. A study with $N = 21$ teachers and $N = 308$ students in Grade 2 and Grade 4 further stresses the importance of teachers in teaching metacognitive abilities (van Loon et al., 2021). Teachers were observed and audio-recordings were made as they provided instruction on a predefined task. Teachers' instructions were evaluated with regard to whether they instructed children on cognitive and metacognitive strategies and whether they used teacher-directed (i.e., directly instructing children how to work on a task) or student-centered (i.e., giving children autonomy to regulate their own learning) instructions. The results indicated that student-centered instructions positively predicted students' restudy and self-monitoring, suggesting that not only the content of the instructions itself, but also the way they were given affected children's metacognition. Finally, a study with $N = 587$ tenth-graders from Uganda demonstrated that the relationship between perceived cognitive activation in the classroom and metacognitive ability was fully mediated by self-efficacy and enjoyment, suggesting that cognitive activation is at least indirectly associated with metacognitive ability (Ekatushabe et al., 2021). Given the importance of motivating and cognitively activating teaching for the overall learning process (see Cents-Boonstra et al., 2021), these two

teaching qualities might be of relevance when exploring teachers' role in the Dunning-Kruger effect and hence the accuracy and direction of students' judgements of their own ability.

The Dunning-Kruger effect and political knowledge

Political knowledge is usually considered to be a major tool for understanding the political world and everything related to it. A widely accepted definition of political knowledge is that it encompasses "the range of factual information about politics that is stored in long-term memory" (Delli Carpini & Keeter, 1996: p. 10). As such, political knowledge helps people understand political debates and the relevance thereof, to sort and categorize political information, and to become aware of their own political needs and preferences and what political actions and decisions must be pursued to satisfy these needs and preferences (Cramer & Toff, 2017). Political knowledge is therefore an integral component of political competence (Alscher et al., 2023), which permeates everyday life.

In recent years, social media has fuelled the dissemination of political news and complicated the identification of false information (see Wineburg et al., 2022). In addition, the 'echo chamber' phenomenon has produced environments that reinforce existing views and opinions, making it more difficult to detect false facts. However, accurately assessing one's own political knowledge is particularly important. For instance, a Dutch online survey ($N = 545$ eligible voters) about the 2016 EU refugee crisis revealed that individuals at the political extremes were overconfident regarding their knowledge of the refugee crisis (Van Prooijen et al., 2018). Another online survey from the Netherlands with $N = 13,323$ complete responses about a referendum on an EU treaty with Ukraine yielded similar results (Van Prooijen & Krouwel, 2020). Domain-specific and general overconfidence predicted anti-establishment voting. In addition, analyses of data from Latin America have demonstrated that overconfident individuals are more likely to endorse authoritarian measures (Marisguia & Dias, 2021). Data from Spain also suggests that overconfident individuals are more likely to believe in conspiracy theories, such as flat-earthism (Arroyo-Barrigüete et al., 2023). Finally, a study with 3,000 adults from the U.S. has shown that overconfidence is a substantive and statistically significant predictor of ideological extremism, vote abstention and partisan identification (Ortoleva & Snowberg, 2015).

Past research has indicated that the pattern of the Dunning-Kruger effect can also be found in the context of political knowledge. In a US study, $N = 1,047$ and $N = 1,559$ adult online survey respondents completed a political knowledge quiz with five questions and were subsequently asked to provide a self-assessment of their mastery of the knowledge battery (Anson,

2018). The results demonstrated that low performers tended to overestimate, while high performers tended to underestimate their political knowledge. As suggested by the Dunning-Kruger effect, the average degree of overestimation among low performers was greater than the average degree of underestimation among high performers. Moreover, a repeated-measures cross-sectional survey with two representative samples of the Finnish voting-age population in 2008 ($N = 1,020$) and 2020 ($N = 1,097$) assessed political knowledge based on nine items and self-assessed political sophistication using a five-point Likert scale and the question “How much would you say you know about politics and public affairs?” (Rapeli, 2023). The study’s results suggest that overconfidence among low performers was more prevalent in 2020 than in 2008. However, while it is claimed that the study’s results confirm the existence of the Dunning-Kruger effect, important information and analyses are lacking, especially regarding low performers’ and high performers’ judgment accuracy, in order to confirm the existence of the Dunning-Kruger effect.

Present study

With this study, we aim to add to the existing literature on the Dunning-Kruger effect by investigating students’ political knowledge in the high school context. Focusing on high school students might be particularly insightful because the plasticity of metacognitive ability is still high during adolescence and differences between age groups are relatively large. Furthermore, we examined what role teachers play regarding discrepancies in high school students’ objectively measured and subjectively perceived political knowledge. To this end, we formulated the following three research questions and hypotheses:

1. Is the Dunning-Kruger effect (e.g., relatively large overestimation among low performers and relatively small underestimation among high performers) present in Grade 7 and Grade 10 with regard to political knowledge?

We hypothesize that the pattern of findings typical of the Dunning-Kruger effect is present in Grade 7 and Grade 10 with regard to political knowledge. Prior research suggests that the Dunning-Kruger effect can be found for political knowledge, too (Anson, 2018; Rapeli, 2023). While this prior evidence is based on adult samples and employs more less comprehensive measures of political knowledge, we nonetheless assume that we will find the Dunning-Kruger effect among high school students.

2. How does the Dunning-Kruger effect differ between Grade 7 and Grade 10 with regard to political knowledge?

We hypothesize that students in Grade 10 estimate their political knowledge more accurately. The discrepancies between students' objective and subjective political knowledge should thus be smaller in Grade 10 than in Grade 7. The proximate reason for this result would be that tenth-graders have better metacognitive abilities (see Weil et al., 2013) and therefore are better able to assess their political knowledge. This finding would support the notion that systematic differences in metacognitive insight between low and high performers cause the Dunning-Kruger effect (e.g., Hiller, 2023; Kruger & Dunning, 1999; Schlösser et al., 2013). On the contrary, finding no differences between seventh-graders and tenth-graders would give support to the notion that the Dunning-Kruger effect might be a statistical artifact (e.g., Dunkel et al., 2023; Gignac & Zajenkowski, 2020; Gignac & Zajenkowski, 2023).

3. How are students' perceptions of teacher behavior (cognitive activation and motivational quality) associated with judgement accuracy and judgement direction (i.e. overconfidence and underconfidence)?

We hypothesize that perceived cognitive activation is related to judgement accuracy. Prior empirical evidence suggests that perceived cognitive activation and metacognitive ability are indirectly associated with one another (Ekatushabe et al., 2021), suggesting that students who perceive their teacher as cognitively activating might be able to assess their political knowledge more accurately.

In addition, we hypothesize that perceived motivational quality is related to judgment accuracy, but also overconfidence. Prior research suggests that motivating and engaging teaching is an important determinant of successful learning (Cents-Boonstra et al., 2021), including in civic education (Bayram-Özdemir et al., 2016). However, research regarding the relationship between motivational quality and metacognition as well as the Dunning-Kruger effect is scarce (cf. Sypré et al., 2022). We assume that the importance of motivating teaching for general learning is transferable to learning metacognitive skills, which is why students who receive motivating instruction should be able to judge their ability more accurately. Furthermore, we assume that students who perceive their teachers as particularly motivating will be more motivated and, as a consequence of higher self-expectations, more prone to overestimation (see Blanton et al., 2001).

Method

Transparency and openness

Below, we report how we determined our sample size, all measures in the study, and how we treated missing data. The analysis code and additional research material are available at

https://osf.io/zvfes/?view_only=8556f97ccba14d4787204cea16237d10. Data for this study will be shared upon request. All analysis were carried using R, version 4.1.0, and the “stats” (R Core Team, 2021) and “lavaan” (Rosseel, 2012) packages. This study’s design and analysis were not pre-registered.

Participants and design

Participants were 1,047 students attending Grade 7 ($n = 613$, $M_{\text{age}} = 12.47$; 47.4% female; 44.4% with an immigrant background) or Grade 10 ($n = 434$, $M_{\text{age}} = 15.49$; 47.7% female; 35.9% with an immigrant background). The participating students came from 19 different schools and 69 different classes. Data were collected on site in each class separately. In Germany, upon completing elementary school, typically around the age of 10, students are sorted into distinct secondary education tracks. This allocation process primarily hinges on students' achievement and aptitude. While substantial regional variability in Germany's educational landscape exists, the prevalent model is the "tripartite" system comprising *Hauptschule*, *Realschule*, and *Gymnasium*. Among the three, *Hauptschule* is most strongly oriented towards vocational training, *Realschule* serves as an intermediary option, and *Gymnasium* prepares students for university entrance. For the sake of clarity, this study distinguishes between schools that prepare students for vocational education and training (*Hauptschule* and *Realschule*), and *Gymnasium*, which readies students for university admission. We sampled schools from rural, semi-urban and urban areas of Western Germany in the federal state of North Rhine-Westphalia. The student response rate was 72%. Students not reached either did not have informed parental consent or were absent on the test day due to illness or quarantine measures. 41.8% of participating students had an immigrant background (operationalized by language spoken at home), which is very close to the official percentage of students with an immigrant background in the same federal state in the school year 2021/22 (42.7%; operationalization not specified; IT.NRW, 2022). The study was conducted in compliance with the German Research Foundation's (*Deutsche Forschungsgemeinschaft*, DFG) guidelines for good scientific practice. In addition, the university's Joint Ethics Committee examined and approved the research project. Students could only participate in the study if they presented an informed consent signed by their parents at the beginning of the assessment. Without informed consent, students had to leave the classroom. Each class was awarded 50 EUR for their participation.

Procedure

Data were collected in fall and winter 2021/2022 in the context of the *Study on the development of political and social competence in adolescence* (German: *Studie zur Entwicklung politischer und gesellschaftlicher Kompetenz im Jugendalter*, EPKO). Data collection took place in the morning and lasted about two and half hours. After distributing laptop computers to the students, trained research assistants explained the study procedure to the students. The study consisted of four parts: (1) a 60-minute political knowledge test, (2) an 18-minute cognitive ability test, (3) a six-minute reading comprehension test, and (4) a 40-minute questionnaire. After each section, students were given short breaks.

Measures

Objective political knowledge

Political knowledge was measured using a comprehensive test instrument that matched the abilities of students in Grade 7 and Grade 10 (Alscher et al., 2022). In total, the instrument consisted of 99 items. The items were evenly distributed among three content areas (polity, politics and policy) and three sub-skills (identify and describe, explain and analyze, and argue and deliberate). Each item was a forced-choice question with three incorrect and one correct answer option. For each item, a correct response was coded as 1 and an incorrect response was coded as 0. Seventh-graders answered 63 and tenth-graders 70 items. 34 items were answered by students from both grade levels. While most items were newly developed, some items were adopted from another study of political knowledge (Goll et al., 2011). In order to link the test results of students in Grade 7 and Grade 10, we calculated a unidimensional multigroup 1-parameter model using R and the “TAM” package (Robitzsch et al., 2021). Before calibration, we detected rapid guessing behavior, which led to the deletion of 6.76% of the answers. We set the threshold for deletion at a processing time of seven seconds or less. This threshold represents the local minimum between the two modes in the response time distribution (Wise, 2017). For every participant, we estimated a person parameter (θ). The overall WLE reliability was .89 (Grade 7: .89; Grade 10: .88).

Subjective political knowledge

Subjective political knowledge was measured with a three-item scale. Two items were drawn from the Political Efficacy Kurzskala (PEKS, Beierlein et al., 2014), and one item was newly developed. The introductory question read as follows: “To what extent do the following statements apply to you?” The three items read (a) “In general, I know a lot about politics”, (b) “I

am good at understanding and assessing important political issues”, and (c) “I have the confidence to take active part in a discussion about political issues.” The students were asked to rate each item on a four-point scale ranging from 1 (= “strongly disagree”) to 4 (= “strongly agree”). The scale’s overall alpha reliability was 0.80 (0.79 and 0.81 for students in Grade 7 and Grade 10, respectively).

Judgement accuracy

Judgement accuracy was operationalized as the absolute value of the difference between subjective and objective political knowledge, i.e., |subjective political knowledge – objective political knowledge|. For this purpose, objective and subjective political knowledge were z-standardized for each grade level (Grade 7 and Grade 10) separately.

Judgement direction

Judgement direction was operationalized as the dichotomized difference between subjective and objective political knowledge. A value above 0 indicated overconfidence and was recoded as 1, and a value below 0 indicated lack of confidence and was recoded as 0. Again, objective and subjective political knowledge were z-standardized beforehand for each grade level (Grade 7 and Grade 10) separately.

Perceived cognitive activation

Perceived cognitive activation in civic education was measured with a four-item scale. Specifically, the students were asked to what extent their teacher establishes a connection to everyday life and creates links with students' prior knowledge. The items were adapted from the Progress in International Reading Literacy Study 2016 (PIRLS, Hußmann et al., 2020). An example item reads “Our teacher makes the subject matter clear to us with examples from real politics.” Students were asked to rate each item on a four-point scale ranging from 1 (= “strongly disagree”) to 4 (= “strongly agree”). The scale’s overall alpha reliability was 0.86 (0.87 and 0.89 for students in Grade 7 and Grade 10, respectively).

Perceived motivational quality

Perceived motivational quality in civic education was measured with a five-item scale. The items were adapted from the *Education through Language and Writing* study (German: *Bildung durch Sprache und Schrift*, BiSS, see McElvany et al., 2019). An example item reads “Our teacher gets us excited about political issues.” Again, students were asked to rate each item on a four-point scale ranging from 1 (= “strongly disagree”) to 4 (= “strongly agree”). The scale’s

overall alpha reliability was 0.91 (0.92 and 0.90 for students in Grade 7 and Grade 10, respectively).

Control Variables

Cultural capital was operationalized by the reported number of books in the student's home (adapted from PISA 2015, see Mang et al., 2019). Responses ranged from 1 (= "0 to 10 books") to 6 (= "more than 500 books"). For immigrant background, students were asked which language they speak at home (adapted from IGLU 2016, see Hußmann et al., 2020). For those students who reported speaking a language other than German at home at least sometimes, an immigrant background was assumed. Furthermore, Grade 7 was coded as 0 and Grade 10 as 1. In terms of school type, schools preparing students for vocational education were coded as 0 and schools preparing students for university entry were coded as 1. For school achievement, students' self-reported Math and German grades ranging from 1 (= "inadequate") to 6 (= "very good") were used. Regarding students' sex, male was coded as 0 and female as 1.

Data Analysis

With regard to the first research question, we visually inspected the relationship between objective and subjective political knowledge for each grade level separately. In addition, we ran one-way ANOVAs to include a statistically more robust method of analyzing the Dunning-Kruger effect (see Gignac & Zajenkowski, 2020). The one-way ANOVAs were used to compare differences in students' judgement accuracy across the performance spectrum (quartiles) within each grade. For the second research question, we visually compared the relationships between objective and subjective political knowledge in Grade 7 and 10. Moreover, we ran a two-way ANOVA in which we compared differences in judgement accuracy among students of a given performance group (quartile) in one grade level with the corresponding performance group in the other grade level. Furthermore, we tested for heteroscedasticity in the relationship between objective and subjective political knowledge using the Breusch-Pagan test in each grade level separately. Prior to all analyses in the context of research question 1 and 2, we z-standardized the scores for objective and subjective political knowledge for each grade level separately.

For the third research question, two structural equation models using judgement accuracy and judgement direction as dependent variables were analyzed. We used the full information maximum likelihood (FIML) method to estimate the models with missing data. We evaluated global model fit with the comparative fit index (CFI), the Tucker-Lewis Index (TLI), the root-mean-square error of approximation (RMSEA), and the standardized root mean square

residual (SRMR) using the following fit criteria for acceptable fit: $CFI \geq .900$, $TLI \geq .900$, $RMSEA \leq .100$, $SRMR < .100$ (Weston & Gore, 2006). We controlled for the clustered structure of the data using class ($n = 69$) as the cluster variable, thus enabling the computation of robust standard errors. For the sake of representativeness, the data were weighted according to the number of students attending each school track in North Rhine-Westphalia. Weighting by immigrant background or gender was not necessary because the distribution in the data was almost identical to the distribution in the base population (i.e., students in North Rhine-Westphalia attending general education schools). Prior to the analyses in the context of research question 3, we z-standardized the scores for objective and subjective political knowledge across the two grade levels. Grade level was included in the analysis as a control variable.

Results

Descriptive Statistics

Prior to standardization, we examined whether objective political knowledge was normally distributed. We performed a Kolmogorov-Smirnov test, which showed that the distribution did not depart significantly from normality ($D = 0.04$, $p = .15$). Since subjective political knowledge is not interval scaled, a statistic test of normality would not yield meaningful results. However, a visual inspection also suggested that the data are most likely normally distributed (see figure A in the appendix). The mean score for objective political knowledge was slightly above 0. This is because during the scaling process, the average score of seventh-graders was set to 0. The mean score for tenth-graders was 0.53. The mean scores for subjective political knowledge, motivational quality, and cognitive activation were around the theoretical midpoint of the scales. The mean score for judgment direction ($M = 0.48$) indicates that approximately one half of participants were overconfident and the other half experienced a lack of confidence. Mean scores for grades in Math and German were similar to each other ($M = 4.23$ and 4.33 , respectively). The mean score $M = 3.09$ for cultural capital (i.e., number of books in the students' home) was a bit below the theoretical midpoint and slightly higher than in comparable German studies (e.g., Heyder et al., 2021; $M = 2.88$).

The bivariate correlations in Table 1 show that objective and subjective political knowledge were moderately correlated ($r = .34$). This is in line with the findings described above for the relationship between objective and subjective ability (e.g., Zell & Krizan, 2014). Judgement accuracy was negatively correlated with subjective political knowledge ($r = -.06$, $p = .04$). Otherwise, judgement accuracy had no statistically significant correlations with any

other variable. Judgement direction (i.e., overconfidence) was negatively correlated with objective political knowledge ($r = -.44, p < .00$) and positively correlated with subjective political knowledge ($r = .47, p < .00$). Furthermore, judgement direction was positively correlated with motivational quality and immigrant background and negatively correlated with enrollment in the university-prep track, Grade 10 as opposed to Grade 7, German grade and cultural capital.

Table 1

Means, standard deviations, and correlations

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. ^{ab} Objective political knowledge	–												
2. ^a Subjective political knowledge	.34	–											
3. ^{bc} Judgement accuracy	-.03	-.06	–										
4. ^c Judgement direction (1 = overconfidence)	-.44	.47	-.02	–									
5. Motivational quality	.13	.34	.02	.16	–								
6. Cognitive activation	.28	.31	-.03	.04	.64	–							
7. School type (1 = academic track)	.27	.14	-.06	-.08	.04	.07	–						
8. Grade level (1 = Grade 10)	.30	.05	.01	-.18	.08	.22	-.16	–					
9. Math grade	.27	.19	-.01	-.06	.09	.08	.17	.01	–				
10. German grade	.32	.21	-.01	-.08	.09	.11	.09	.06	.58	–			
11. Cultural capital	.35	.19	.00	-.10	.05	.08	.23	-.02	.17	.20	–		
12. Immigrant background	-.25	-.10	-.04	.13	-.02	-.03	.05	-.06	-.09	-.15	-.18	–	
13. Sex (1 = female)	-.03	-.12	.05	-.05	-.01	.04	.02	.02	-.05	.14	.01	.07	–
<i>M</i>	0.23	2.32	3.26	0.48	2.30	2.54	0.43	0.41	4.23	4.33	3.09	0.42	0.48
<i>SD</i>	0.86	0.72	0.69	0.50	0.79	0.79	0.50	0.49	1.10	0.98	1.49	0.49	0.50
Actual Range	-1.7 – 3.1	1–4	0 – 4.2	0–1	1–4	1–4	0–1	0–1	1–6	1–6	1–6	0–1	0–1
<i>N</i> _{items}	99	3	–	–	5	4	1	1	1	1	1	1	1
<i>N</i> _{responses}	1047	986	986	986	930	923	1047	1047	992	992	1018	1022	1024
Cronbach's α	^c .89	.80	–	–	.91	.86	–	–	–	–	–	–	–
Grade 7	^c .89	.79	–	–	.92	.87	–	–	–	–	–	–	–
Grade 10	^c .88	.81	–	–	.90	.89	–	–	–	–	–	–	–

Note. *M* and *SD* represent mean and standard deviation, respectively.

^aScores were z-standardized for all subsequent analyses. ^bActual range rounded to one decimal place. ^cScores are based on z-standardized values of objective political knowledge and subjective political knowledge.

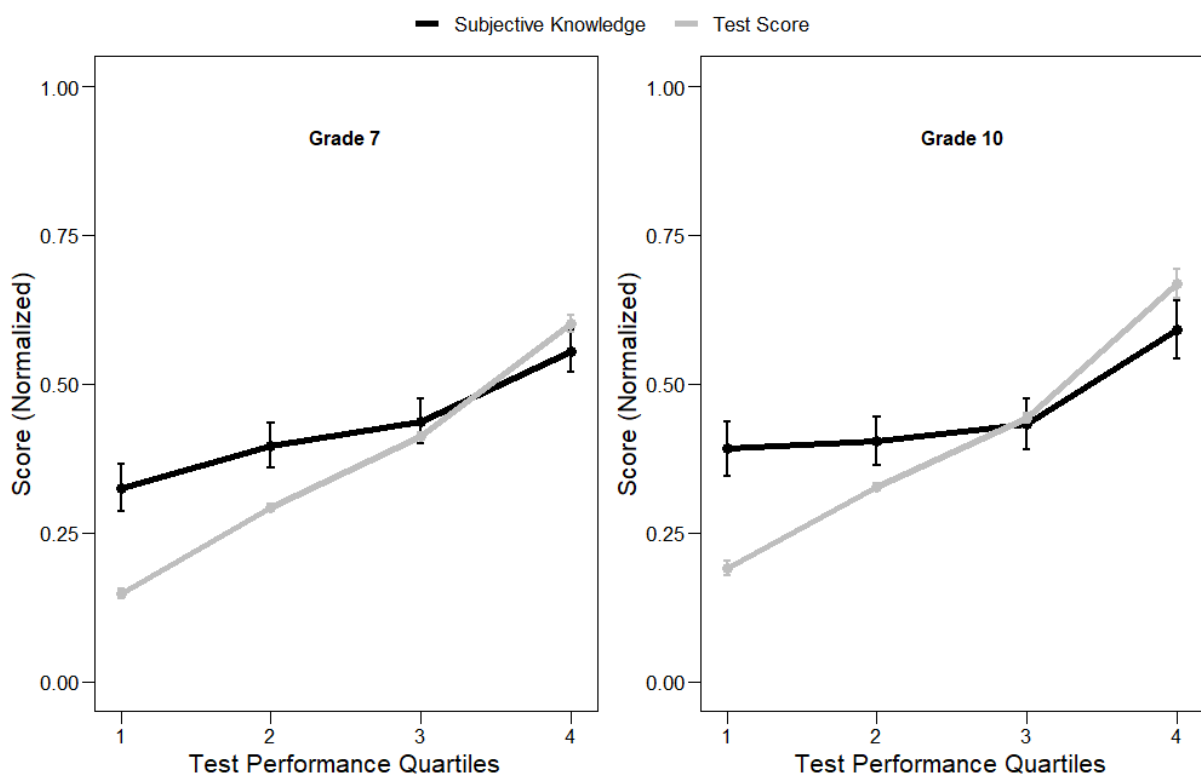
Correlations printed in bold reached statistical significance at the 5% level.

The relationship between objective and subjective political knowledge in grades 7 and 10

A visual inspection of the relationship between objective and subjective political knowledge indicates that the pattern typical for the Dunning-Kruger effect exists in Grade 7 and Grade 10. In both grade levels, low performers tended to overestimate their political knowledge and high performers tended to underestimate their political knowledge, with low performers having on average less accurate estimates than high performers. Furthermore, in both grade levels, individuals in the third performance quartile seemed to provide the most accurate estimates of their political knowledge.

Figure 1

The Dunning-Kruger effect in Grade 7 and Grade 10



Note. Subjective political knowledge and test score were first z-standardized and then rescaled to range from 0 to 1.

The average differences between objective and subjective political knowledge (i.e., judgment accuracy) ranged from $|\Delta| = .03$ to $|\Delta| = .18$ within performance quartiles in Grade 7 and from $|\Delta| = .01$ to $|\Delta| = .20$ within performance quartiles in Grade 10 (see Figure 1). In both grade levels, the average judgement accuracy was the lowest in the first quartile ($|\Delta| = .18$ and $.20$ for Grade 7 and Grade 10, respectively) and highest in the third quartile ($|\Delta| = .03$ and $.01$

for Grade 7 and Grade 10, respectively). One-way ANOVAs for each grade level demonstrate that judgement accuracy was statistically significantly different across quartiles in Grade 7 ($F(1, 571) = 75.81, p < .01$) and Grade 10 ($F(1, 411) = 83.99, p < .01$). Subsequent pairwise t -tests illustrate that all differences in average judgement accuracy between quartiles were statistically significant in both grade levels (see table 2). Differences in average differences across quartiles range from $\Delta = .07$ to $\Delta = .23$ in Grade 7 and from $\Delta = .07$ to $\Delta = .28$ in Grade 10. In both grade levels, the differences increase as the quartiles become further apart. The Breusch-Pagan test adds to this impression, as it fails to reject the null hypothesis of homoscedasticity for both grade levels (Grade 7: $\chi^2(1) = 0.78, p = .38$; Grade 10: $\chi^2(1) = 0.58, p = .45$).

Table 2

Pairwise t-tests of average differences across quartiles within and across grade levels

Grade level	Quartiles	Grade 7				Grade 10			
		1	2	3	4	1	2	3	4
Grade 7	1	–							
	2	0.08	–						
	3	0.15	0.07	–					
	4	0.23	0.15	0.08	–				
Grade 10	1	0.02				–			
	2		0.03			0.12	–		
	3			0.04		0.21	0.09	–	
	4				0.03	0.28	0.16	0.07	–

Note. Grade 7: $F(1, 571) = 75.81, p < .01$; Grade 10: $F(1, 411) = 83.99, p < .01$; Grade 7/10: $F_Q(3, 981) = 52.51, p_Q < .01, F_G(1, 981) = 1.53, p_G = .22$.

Difference values printed in bold reached statistical significance at the 5% level.

With regard to the second research question, the graphs illustrate that the results are very similar in the two grade levels. A two-way ANOVA comparing average judgement accuracy across performance quartiles and grade level reaffirms that judgement accuracy was statistically significantly different across quartiles ($F(3, 981) = 52.12, p < .01$), and also shows that judgement accuracy did not statistically significantly differ between grade levels ($F(1, 981) = 1.53, p = < .22$). A comparison of the average judgement accuracy of corresponding quartiles between the two grade levels shows that the differences ranged from $\Delta = .02$ to $\Delta = .04$. None of the

differences across grade levels reached statistical significance, indicating that even when controlling for each quartile separately, no statistically significant differences exist across the two grade levels with regard to judgement accuracy.

Perceptions of teacher behavior and judgement accuracy and direction

To answer the third research question, two structural equation models were fitted using judgement accuracy and judgement direction as dependent variables. All model fit indices were within the recommended values, indicating good fit for both models (see Table 3). In accordance with our hypothesis, perceived motivational quality was positively associated with judgement accuracy ($\beta = .16, p = .01$) and judgement direction ($\beta = .22, p < .01$; i.e., motivational quality was associated with overestimation).

Table 3

Results of the structural equation models

	Judgement accuracy		Judgement direction	
	β	p	β	p
Motivational quality	.16	.01	.22	<.01
Cognitive activation	-.17	.02	-.03	.72
School type (1 = academic track)	-.07	.11	-.10	.01
Grade level (1 = Grade 10)	.02	.49	-.20	<.01
Math grade	.03	.45	-.02	.66
German grade	-.01	.80	-.05	.23
Cultural capital	.01	.73	-.05	.10
Immigrant background	-.04	.42	.12	<.01
Sex (1 = female)	.07	.05	-.04	.28
R^2	.03		.12	
CFI	.962		.961	
TLI	.955		.953	
RMSEA	.041		.042	
SRMR	.055		.056	

Note. $N = 961$. Estimates printed in bold reached statistical significance at the 5% level.

Rather counterintuitively, perceived cognitive activation was negatively associated with judgement accuracy ($\beta = -.17, p = .02$). Furthermore, cognitive activation was not associated with judgement direction ($\beta = -.03, p < .72$). Among the control variables, school track and grade level were associated with underestimation ($\beta = -.10, p = .01$ and $\beta = -.20, p < .01$,

respectively). In addition, immigrant background was associated with overestimation ($\beta = .12$, $p < .01$). Students' sex, grades in Math and German as well as cultural capital were not statistically significantly associated with either judgement accuracy or judgement direction.

Discussion

In the present study, we aimed to take a novel approach to investigating the Dunning-Kruger effect by analyzing and comparing it in two different grade levels. We found the pattern typical of the Dunning-Kruger effect for students' political knowledge in Grade 7 and Grade 10. Low performers estimated their ability less accurately than high performers. Students in the third performance quartile provided the most accurate estimates, which, again, is similar to what Kruger and Dunning (1999) found in their original study. Empirical analyses of students' judgement accuracy reaffirm these findings and suggest that the Dunning-Kruger effect is present in our data.

With regard to potential differences between Grade 7 and Grade 10, we found that visually, no differences in the pattern typical of the Dunning-Kruger effect could be discerned between the two grade levels. Similarly, statistically, no significant differences between seventh-graders' and tenth-graders' average judgment accuracy could be detected. Given the premise that the level of metacognitive abilities should differ between seventh-graders and tenth-graders, we conclude that our results at least do not support the assumption that metacognitive abilities are responsible for the Dunning-Kruger effect in our data.

The fact that no differences in the Dunning-Kruger effect were found for students in Grade 7 and Grade 10 could either mean that the average metacognitive abilities of seventh-graders and tenth-graders do not differ, that metacognitive abilities are not responsible for the Dunning-Kruger effect, or that an important but unknown third variable differs systematically between the two grade levels, biasing the influence of metacognitive abilities in this context. An unknown systematically differing third variable seems rather unlikely given the sample size and structure. In addition, no absolute differences in the metacognitive ability of students in the two grade levels seems unlikely in light of the existing literature (e.g., Schneider et al., 2022). However, it is conceivable that the Dunning-Kruger effect is driven not so much by the absolute level of metacognitive ability but by the relative level of metacognitive ability (i.e., students' metacognitive ability compared to their peers' metacognitive ability). Another option, of course, would be that metacognitive abilities are simply not responsible for the Dunning-Kruger effect and that the effect is merely a statistical artifact driven by the better-than-average effect and regression-to-the-mean, as argued by critics (e.g., Gignac & Zajenkowski, 2023).

However, what speaks against the better-than-average effect is that the Dunning-Kruger effect is not only found when individuals are tasked with judging their performance relative to others, but also when tasked with judging their absolute performance (see Jansen et al., 2021). Accordingly, conclusive clarification of the role of metacognitive abilities cannot be provided at this point and future research should strive to develop further novel approaches that overcome known pitfalls for testing the Dunning-Kruger effect and its underlying mechanisms.

We furthermore investigated what role teacher behavior plays for judgment accuracy and judgment direction. We found that perceived motivational quality is related to judgement accuracy, but also to overconfidence. Cognitive activation is negatively related to judgment accuracy and not related to over- or underconfidence.

The findings for perceived motivational quality are consistent with our hypotheses. In general, motivational quality is associated with more accurate estimates. However, students who perceive their teacher as particularly motivating tend to overestimate their political knowledge (see figure B in the appendix). The negative relationship between cognitive activation and judgment accuracy is rather counterintuitive. Closer examination of the bivariate relationships between perceived cognitive activation and objective political knowledge as well as subjective political knowledge suggests that the negative relationship between cognitive activation and judgment accuracy is mainly driven by students who perceive their civics courses to be very cognitively activating and these students' tendency to overestimate their political knowledge (see figure C in the appendix). Future research should try to investigate why the positive relationship between cognitive activation and objective political knowledge does not hold for students who perceive their civics courses as particularly cognitively activating.

In addition to the strong data source, strengths of this study include the innovative approach of using two cohorts at different developmental stages to study the Dunning-Kruger effect and linking teacher behavior to judgment accuracy and judgment direction. Despite its strengths, this study also comes with a few limitations. When investigating the Dunning-Kruger effect, study participants are usually asked to estimate the number of questions they answered correctly or their relative performance compared to their peers. In this study, we asked the participating students in a more abstract way about their subjective political knowledge. However, judging by the graphical illustrations of this study's findings, it appears that the slightly different assessment of subjective political knowledge does not alter the general pattern of findings of the Dunning-Kruger effect. In addition, of course, the study would have benefitted from a measure of students' metacognitive abilities. However, metacognition is a fuzzy concept and its

measurement might be confounded by response bias (Fleming & Lau, 2014) or optimism bias (Krueger & Mueller, 2002).

The results bear important implications for educational research and practice. In recent years, the number of studies scrutinizing the metacognitive interpretation of the Dunning-Kruger effect seems to have increased (e.g., Gignac & Zajenkowski, 2023; Magnus & Peresetsky, 2022). Nevertheless, several studies continue to defend the metacognitive explanation of the Dunning-Kruger effect (e.g., Hiller, 2023) and some studies even disregard alternative explanations such as the better-than-average effect and regression-to-the-mean (e.g., West & Eaton, 2019; Zhou & Jenkins, 2020).

Educational aspects such as teacher behavior have so far played only a limited role in the investigation of the Dunning-Kruger effect (cf. van der Beek et al., 2017). However, the Dunning-Kruger effect is also an important resource for educational research, especially when it comes to objective and self-assessed ability and the discrepancy between these two measures. Overestimation and underestimation can affect educational outcomes but also political attitudes and might exacerbate conflict (e.g., Dunlosky & Rawson, 2012; Van Prooijen et al., 2018). Research results on teaching metacognitive skills give rise to the hope that schools and teachers can help students come to more accurate assessments of their own abilities (e.g., Dignath & Büttner, 2008). This study has shown that the relationship between teacher behavior and judgment accuracy as well as judgment direction is complex and that more research is needed to reach a better understanding of this relationship.

Conclusion

To better understand the role that metacognitive abilities play in the Dunning-Kruger effect, the current study used a novel approach and data set with participants from two different grade levels. The results suggest that the typical pattern of the Dunning-Kruger effect can be found in both grade levels and does not seem to be affected by any systematic differences in metacognitive abilities across grade levels. However, this study does not allow for a conclusive assessment of the role of metacognitive abilities in the Dunning-Kruger effect. Furthermore, we explored the relationship between teacher behavior and judgement accuracy as well as judgment direction. The results suggest that teacher behavior plays an important role for students' self-assessment. Additional research efforts are needed to enhance our understanding of the impact of teaching behavior on self-assessments by children and adolescents. This study joins the ranks of research that have made at least an incremental contribution to understanding the Dunning-Kruger effect.

References

- Alscher, P., Grecu, A. L., Ludewig, U., & McElvany, N. (2023). Towards the measurability of a competence model for school-based civic education. In N. McElvany, A. L. Grecu, R. Lorenz, M. Becker, C. Dignath, H. Gaspard, & F. Lauer mann (Eds.), *50 Jahre Schulentwicklung – Leitthemen der empirischen Bildungsforschung* (pp. 230–261). Beltz Juventa. <https://doi.org/10.17605/OSF.IO/TR8JP>
- Alscher, P., Ludewig, U., & McElvany, N. (2022). Civic Literacy – zur Theorie und Messbarkeit eines Kompetenzmodells für die schulische politische Bildung. *Zeitschrift Für Erziehungswissenschaft*, 25(5), 1221–1241. <https://doi.org/10.1007/s11618-022-01085-0>
- Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, 13(2), 57. <https://doi.org/10.1037/1076-898x.13.2.57>
- Anson, I. G. (2018). Partisanship, political knowledge, and the Dunning-Kruger effect. *Political Psychology*, 39(5), 1173–1192. <https://doi.org/10.1111/pops.12490>
- Arroyo-Barrigüete, J. L., Bellón Núñez-Mera, C., Labrador Fernández, J., & Nicolas, V. L. de (2023). Dunning–Kruger effect and flat-earthers: An exploratory analysis. *Public Understanding of Science*. <https://doi.org/10.1177/09636625231166255>
- Artelt, C., Neuenhaus, N., Lingel, K., & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, 63(1), 18–25. <https://doi.org/10.1026/0033-3042/a000106>
- Baars, M., Leopold, C., & Paas, F. (2018). Self-explaining steps in problem-solving tasks to improve self-regulation in secondary education. *Journal of Educational Psychology*, 110(4), 578–595. <https://doi.org/10.1037/edu0000223>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)
- Bayram-Özdemir, S., Stattin, H., & Özdemir, M. (2016). Youth’s initiations of civic and political discussions in class: Do youth’s perceptions of teachers’ behaviors matter and why? *Journal of Youth and Adolescence*, 45(11), 2233–2245. <https://doi.org/10.1007/s10964-016-0525-z>
- Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2014). *Political Efficacy Kurzskala (PEKS)* (Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)). <https://doi.org/10.6102/ZIS34>

- Blanton, H., Pelham, B. W., DeHart, T., & Carvallo, M. (2001). Overconfidence as Dissonance Reduction. *Journal of Experimental Social Psychology, 37*(5), 373–385. <https://doi.org/10.1006/jesp.2000.1458>
- Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research, 50*(4), 221–229. <https://doi.org/10.1016/j.ijer.2011.08.003>
- Branigan, H. E., & Donaldson, D. I. (2020). Teachers matter for metacognition: Facilitating metacognition in the primary school through teacher-pupil interactions. *Thinking Skills and Creativity, 38*, 1–14. <https://doi.org/10.1016/j.tsc.2020.100718>
- Cents-Boonstra, M., Lichtwarck-Aschoff, A., Denessen, E., Aelterman, N., & Haerens, L. (2021). Fostering student engagement with motivating teaching: an observation study of teacher and student behaviours. *Research Papers in Education, 36*(6), 754–779. <https://doi.org/10.1080/02671522.2020.1767184>
- Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience, 1*(3), 165–174. <https://doi.org/10.1093/scan/ns1024>
- Cramer, K. J., & Toff, B. (2017). The Fact of Experience: Rethinking Political Knowledge and Civic Competence. *Perspectives on Politics, 15*(3), 754–770. <https://doi.org/10.1017/S1537592717000949>
- Davis, S. D., & Chan, J. (2023). Effortful Tests and Repeated Metacognitive Judgments Enhance Future Learning. *Educational Psychology Review, 35*(3). <https://doi.org/10.1007/s10648-023-09803-8>
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. <https://doi.org/10.1007/s11409-008-9029-x>
- Downing, K., Kwong, T., Chan, S.-W., Lam, T.-F., & Downing, W.-K. (2009). Problem-based learning and the development of metacognition. *Higher Education, 57*, 609–621. <https://doi.org/10.1007/s10734-008-9165-x>
- Dunkel, C. S., Nedelec, J., & van der Linden, D. (2023). Reevaluating the Dunning-Kruger effect: A response to and replication of Gignac and Zajenkowski (2020). *Intelligence, 96*. <https://doi.org/10.1016/j.intell.2022.101717>

- Dunlosky, J., Mueller, M. L., Morehead, K., Tauber, S. K., Thiede, K. W., & Metcalfe, J. (2021). Why Does Excellent Monitoring Accuracy Not Always Produce Gains in Memory Performance? *Zeitschrift Für Psychologie*, 229(2), 104–119. <https://doi.org/10.1027/2151-2604/a000441>
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83–87.
- Dunning, D. (2011). Chapter five - The Dunning–Kruger Effect: On Being Ignorant of One's Own Ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 247–296). Academic Press. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105(1), 98–121. <https://doi.org/10.1016/j.obhdp.2007.05.002>
- Ekatushabe, M., Nsanganwimana, F., Muwonge, C. M., & Ssenyonga, J. (2021). The Relationship Between Cognitive Activation, Self-Efficacy, Achievement Emotions and (Meta)cognitive Learning Strategies Among Ugandan Biology Learners. *African Journal of Research in Mathematics, Science and Technology Education*, 25(3), 247–258. <https://doi.org/10.1080/18117295.2021.2018867>
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist*, 34, 906–911. <https://psycnet.apa.org/doi/10.1037/0003-066X.34.10.906>
- Flavell, J. H., Frances L. Green, & Eleanor R. Flavell (1998). The mind has a mind of its own: Developing knowledge about mental uncontrollability. *Cognitive Development*, 13(1), 127–138. [https://doi.org/10.1016/S0885-2014\(98\)90024-7](https://doi.org/10.1016/S0885-2014(98)90024-7)
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8, 443. <https://doi.org/10.3389/fnhum.2014.00443>
- Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, 138(2), 296–321. <https://doi.org/10.1037/a0026556>

- Galambos, N. L., MacDonald, S. W. S., Naphtali, C., Cohen, A.-L., & Frias, C. M. de (2005). Cognitive performance differentiates selected aspects of psychosocial maturity in adolescence. *Developmental Neuropsychology*, 28(1), 473–492. https://doi.org/10.1207/s15326942dn2801_2
- Gignac, G. E., & Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data. *Intelligence*, 80. <https://doi.org/10.1016/j.intell.2020.101449>
- Gignac, G. E., & Zajenkowski, M. (2023). Still no Dunning-Kruger effect: A reply to Hiller. *Intelligence*, 97. <https://doi.org/10.1016/j.intell.2023.101733>
- Goll, T., Richter, D., & Weißeno, G. (2011). Politisches Wissen von Schüler/-innen mit und ohne Migrationshintergrund (POWIS): Ergebnisse einer Studie. In D. Lange (Ed.), *Schriftenreihe der DVPB, Wochenschau Wissenschaft, Entgrenzungen: Gesellschaftlicher Wandel und politische Bildung* (pp. 126–131). Wochenschau Verlag.
- Heyder, A., Weidinger, A. F., & Steinmayr, R. (2021). Only a Burden for Females in Math? Gender and Domain Differences in the Relation Between Adolescents' Fixed Mindsets and Motivation. *Journal of Youth and Adolescence*, 50(1), 177–188. <https://doi.org/10.1007/s10964-020-01345-4>
- Hiller, A. (2023). Comment on Gignac and Zajenkowski, “The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data”. *Intelligence*, 97. <https://doi.org/10.1016/j.intell.2023.101732>
- Hußmann, A., Wendt, H., Bos, W., & Rieser, S. (Eds.). (2020). *IGLU 2016: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Waxmann. <https://elibrary.utb.de/doi/book/10.31244/9783830992936>
- Jansen, R. A., Rafferty, A. N., & Griffiths, T. L. (2021). A rational model of the Dunning–Kruger effect supports insensitivity to evidence in low performers. *Nature Human Behaviour*, 5(6), 756–763. <https://doi.org/10.1038/s41562-021-01057-0>
- Kennedy, J. A., Anderson, C., & Moore, D. A. (2013). When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes*, 122(2), 266–279. <https://doi.org/10.1016/j.obhdp.2013.08.005>
- Koc, E. (2021). Intercultural competence in tourism and hospitality: Self-efficacy beliefs and the Dunning Kruger Effect. *International Journal of Intercultural Relations*, 82, 175–184. <https://doi.org/10.1016/j.ijintrel.2021.04.003>

- Krajč, M., & Ortmann, A. (2008). Are the unskilled really that unaware? An alternative explanation. *Journal of Economic Psychology*, 29(5), 724–738. <https://doi.org/10.1016/j.joep.2007.12.006>
- Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, 82(2), 180–188. <https://psycnet.apa.org/doi/10.1037/0022-3514.82.2.180>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037//0022-3514.77.6.1121>
- Kruger, J., & Dunning, D. (2002). Unskilled and unaware—but why? A reply to Krueger and Mueller (2002). *Journal of Personality and Social Psychology*, 82(2), 189–192. <https://psycnet.apa.org/doi/10.1037/0022-3514.82.2.189>
- Mabe, P., & West, S. (1982). Validity of Self-Evaluation of Ability: A Review and Meta-Analysis. *Journal of Applied Psychology*, 67, 280–296. <https://doi.org/10.1037/0021-9010.67.3.280>
- Magnus, J. R., & Peresetsky, A. A. (2022). A Statistical Explanation of the Dunning-Kruger Effect. *Frontiers in Psychology*, 13, 1–9. <https://doi.org/10.3389/fpsyg.2022.840180>
- Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A., & Reiss, K. (2019). *PISA 2015 Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Waxmann. <https://elibrary.utb.de/doi/book/10.31244/9783830990321>
- Marchant, G. J. (1989). Metateaching: A metaphor for reflective teaching. *Education*, 109(4), 487–489.
- Marisguia, B. A. H., & Dias, A. K. M. (Eds.) (2021). *Overconfidence in Political Knowledge and Its Effects on the Democratic Attitudes of Latin Americans*.
- McElvany, N., Bos, W., Köller, O., Schwabe, F., Ohle-Peters, A. & Asseburg, R. (2019). *Evaluationsvorhaben von Konzepten und Maßnahmen der fachübergreifenden Leseförderung des Primarbereichs im Rahmen der Bund-Länder-Initiative „Bildung durch Sprache und Schrift (BiSS): Schlussbericht*. TU Dortmund.
- McIntosh, R. D., Fowler, E. A., Lyu, T., & Della Sala, S. (2019). Wise up: Clarifying the role of metacognition in the Dunning-Kruger effect. *Journal of Experimental Psychology. General*, 148(11), 1882–1897. <https://doi.org/10.1037/xge0000579>

- McIntosh, R. D., Moore, A. B., Liu, Y., & Della Sala, S. (2022). Skill and self-knowledge: empirical refutation of the dual-burden account of the Dunning–Kruger effect. *Royal Society Open Science*, 9(12), <https://doi.org/10.1098/rsos.191727>
- Metcalfe, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18(3), 159–163. <https://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Miller, T. M., & Geraci, L. (2011). Unskilled but aware: Reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 502–506. <https://doi.org/10.1037/a0021802>
- Murphy, S. C., Barlow, F. K., & Hoppel, W. von (2018). A longitudinal test of three theories of overconfidence. *Social Psychological and Personality Science*, 9(3), 353–363. <https://doi.org/10.1177/1948550617699252>
- Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, 13(2), 179–212. <https://doi.org/10.1007/s11409-018-9183-8>
- Ortoleva, P., & Snowberg, E. (2015). Overconfidence in Political Behavior. *American Economic Review*, 105(2), 504–535. <https://doi.org/10.1257/aer.20130921>
- Perry, J., Lundie, D., & Golder, G. (2019). Metacognition in schools: what does the literature suggest about the effectiveness of teaching metacognition in schools? *Educational Review*, 71(4), 483–500. <https://doi.org/10.1080/00131911.2018.1441127>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rahmani, M. (2020). Medical trainees and the Dunning–Kruger effect: when they don’t know what they don’t know. *Journal of Graduate Medical Education*, 12(5), 532–534. <https://doi.org/10.4300%2FJGME-D-20-00134.1>
- Rapeli, L. (2023). Illusion of knowledge: is the Dunning-Kruger effect in political sophistication more widespread than before? *Journal of Elections, Public Opinion and Parties*, 1–11. <https://doi.org/10.1080/17457289.2023.2214734>
- Robitzsch, A., Kiefer, Thomas, Wu, & Margaret. (2021). *TAM: Test analysis module*. <https://cran.r-project.org/web/packages/TAM/index.html>
- Roebbers, C. M., Schmid, C., & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children’s test performance. *British Journal of Educational Psychology*, 79(4), 749–767. <https://doi.org/10.1348/978185409X429842>

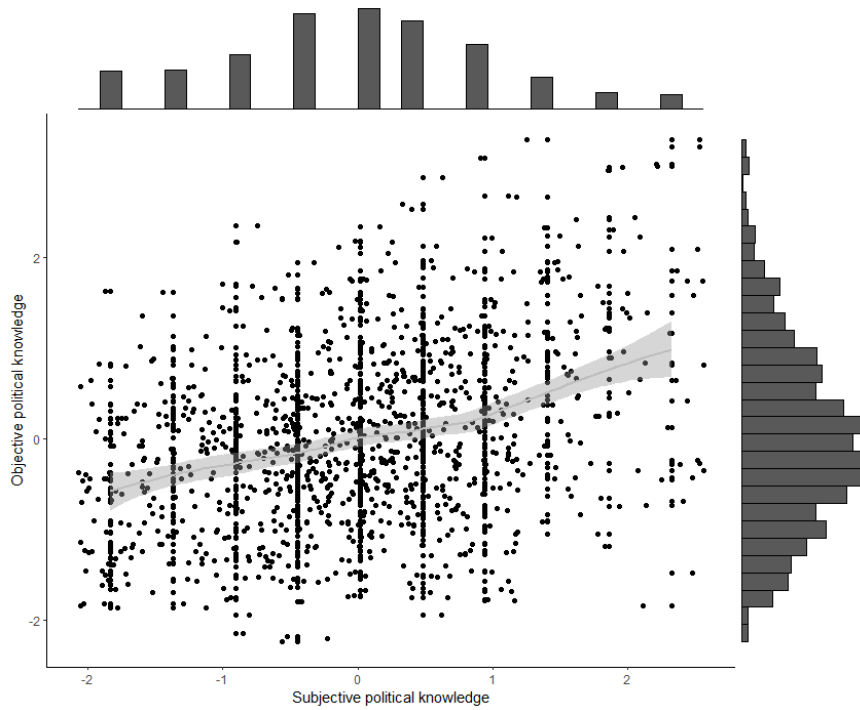
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *2012*, *48*(2), 36. <https://doi.org/10.18637/jss.v048.i02>
- Schlösser, T., Dunning, D., Johnson, K., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the “signal extraction” counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, *39*, 85–100. <https://doi.org/10.1016/j.joep.2013.07.004>
- Schneider, W., Tibken, C., & Richter, T. (2022). The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications. *Advances in Child Development and Behavior*, *63*, 273–307. <https://doi.org/10.1016/bs.acdb.2022.04.006>
- Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2022). Implications of under-confidence and over-confidence in mathematics at secondary school. *International Journal of Educational Research*, *116*. <https://doi.org/10.1016/j.ijer.2022.102085>
- Sullivan, P. J., Ragona, M., & Dithurbide, L. (2019). An investigation into the Dunning–Kruger effect in sport coaching. *International Journal of Sport and Exercise Psychology*, *17*(6), 591–599. <https://doi.org/10.1080/1612197X.2018.1444079>
- Sypré, S., Waterschoot, J., Vansteenkiste, M., Soenens, B., & Verschueren, K. (Eds.) (2022). *Do Teachers Motivate Cognitively Gifted Learners in a Different Way? The Role of Effectiveness, Entity, and Giftedness Beliefs*.
- Syzmanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences*, *21*(5), 493–504. <https://doi.org/10.1016/j.lindif.2011.07.001>
- van der Beek, J. P. J., van der Ven, S. H. G., Kroesbergen, E. H., & Leseman, P. P. M. (2017). Self-concept mediates the relation between achievement and emotions in mathematics. *British Journal of Educational Psychology*, *87*(3), 478–495. <https://doi.org/10.1111/bjep.12160>
- van der Stel, M., & Veenman, M. V. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, *20*(3), 220–224. <https://doi.org/10.1016/j.lindif.2009.11.005>
- van der Stel, M., & Veenman, M. V. J. (2014). Metacognitive skills and intellectual ability of young adolescents: a longitudinal study from a developmental perspective. *European Journal of Psychology of Education*, *29*(1), 117–137. <https://doi.org/10.1007/s10212-013-0190-5>

- van Loon, M. H., Bayard, N. S., Steiner, M., & Roebbers, C. M. (2021). Connecting teachers' classroom instructions with children's metacognition and learning in elementary school. *Metacognition and Learning, 16*(3), 623–650. <https://doi.org/10.1007/s11409-020-09248-2>
- van Prooijen, J.-W. (2021). Overconfidence in radical politics. *The Psychology of Populism, 143–157*.
- van Prooijen, J.-W., & Krouwel, A. P. M. (2020). Overclaiming Knowledge Predicts Anti-establishment Voting. *Social Psychological and Personality Science, 11*(3), 356–363. <https://doi.org/10.1177/1948550619862260>
- van Prooijen, J.-W., Krouwel, A. P. M., & Emmer, J. (2018). Ideological responses to the EU refugee crisis: The left, the right, and the extremes. *Social Psychological and Personality Science, 9*(2), 143–150. <https://doi.org/10.1177/1948550617731501>
- Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., Dolan, R. J., & Blakemore, S.-J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition, 22*(1), 264–271. <https://doi.org/10.1016/j.con-cog.2013.01.004>
- Weston, R., & Gore, P. A. (2006). A Brief Guide to Structural Equation Modeling. *The Counseling Psychologist, 34*(5), 719–751. <https://doi.org/10.1177/0011000006286345>
- Wineburg, S., Breakstone, J., McGrew, S., Smith, M. D., & Ortega, T. (2022). Lateral reading on the open Internet: A district-wide field study in high school government classes. *Journal of Educational Psychology, 114*(5), 893–909. <https://doi.org/10.1037/edu0000740>
- Wise, S. L. (2017). Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications. *Educational Measurement: Issues and Practice, 36*(4), 52–61. <https://doi.org/10.1111/emip.12165>
- Zell, E., & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science, 9*(2), 111–125. <https://doi.org/10.1177/1745691613518075>
- Zell, E., Strickhouser, J., Sedikides, C., & Alicke, M. (2020). The Better-Than-Average Effect in Comparative Self-Evaluation: A Comprehensive Review and Meta-Analysis. *Psychological Bulletin, 146*(2), 118–149. <https://doi.org/10.1037/bul0000218>
- Zhou, X., & Jenkins, R. (2020). Dunning–Kruger effects in face perception. *Cognition, 203*. <https://doi.org/10.1016/j.cognition.2020.104345>

Appendix

Figure A1

Density distributions of z-standardized responses for objective political knowledge and subjective political knowledge

**Figure A2**

Bivariate relationship between motivational quality and the difference between subjective and objective political knowledge

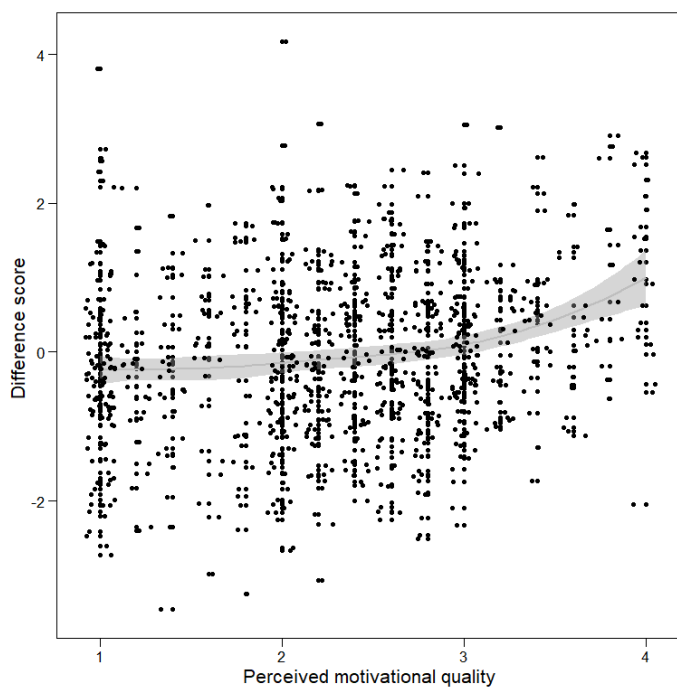
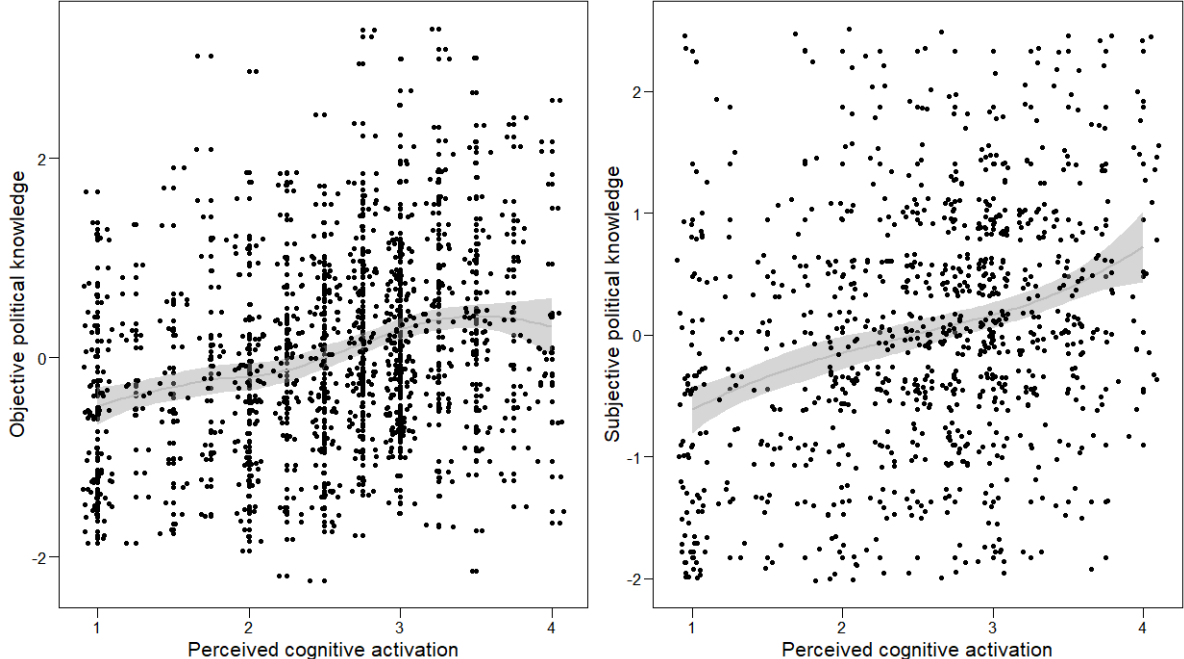


Figure A3

Bivariate relationships between cognitive activation and objective as well as subjective political knowledge



4 Diskussion

Ziel der Dissertation war die Entwicklung und Überprüfung eines Kompetenzmodells, welches die systematische Erfassung von Civic Literacy in der Sekundarstufe I ermöglicht. Hierfür wurde ein Civic Literacy Kompetenzmodell entwickelt und mit Blick auf seine faktorielle Struktur, seine Messinvarianz in der 7. und 10. Klassenstufe sowie der konvergenten und diskriminanten Validität der im Modell enthaltenen Konstrukte überprüft. Der in dem Kompetenzmodell beinhaltete politische Wissenstest wurde mit Blick auf schwierigkeitsgenerierende Merkmale näher betrachtet. Konkret wurde untersucht, inwiefern durchschnittliche Worthäufigkeiten, abgeleitet aus alltäglichen und akademischen Sprachsettings, als Proxy für den Expositionsgrad von in Items behandelten Konzepten mit der Schwierigkeit von Items zusammenhängen. Anschließend wurde untersucht, inwiefern unterschiedliche Qualitätsmerkmale des Politikunterrichts mit verschiedenen Bereichen des entwickelten Kompetenzmodells zusammenhängen. Schließlich wurde der Dunning-Kruger Effekt im Kontext des politischen Wissens untersucht. Konkret wurde untersucht, inwiefern sich der Befund des Dunning-Kruger Effekts in der 7. und 10. Klassenstufe unterscheidet und inwiefern das Lehrkräfteverhalten eine Rolle für den Dunning-Kruger Effekt spielt.

4.1 Diskussion zentraler Ergebnisse

Die Diskussion der zentralen Ergebnisse orientiert sich an den folgenden im Vorfeld festgelegten übergeordneten Forschungsfragen:

- (1) Was umfasst Civic Literacy und wie kann sie bei Jugendlichen gemessen werden? (*Beitrag I, II und III*)
- (2) Inwiefern stehen unterschiedliche Merkmale der schulischen politischen Bildung mit den verschiedenen Bereichen von Civic Literacy in Zusammenhang? (*Beitrag IV*)
- (3) Wie hängen subjektives und objektives politisches Wissen zusammen? (*Beitrag V*)

4.1.1 Was umfasst Civic Literacy und wie kann sie bei Jugendlichen gemessen werden?

Es herrscht Einigkeit darüber, dass Bildung ein wichtiges Instrument ist, um Civic Literacy und mit ihr demokratische Handlungsweisen zu vermitteln (Dewey, 1916; Teegelbeckers et al., 2023). Anders als in anderen Fächern existieren für die politische Bildung in Deutschland jedoch keine verbindlichen nationalen Bildungsstandards. In der Politikdidaktik führte die Einführung der Bildungsstandards in anderen Fächern zu einer breit geführten Diskussion über

Kompetenz in der politischen Bildung (für eine Chronologie siehe Goll, 2019). In den Diskussionen ging es vor allem um die Dimensionen und Facetten, die ein Kompetenzmodell enthalten sollte und weniger um die Frage, ob Bildungsstandards für die politische Bildung wünschenswert seien (Goll, 2019; siehe jedoch auch Sander, 2009). Streitpunkte betrafen (1) die Bezugswissenschaft eines Kompetenzmodells für die politische Bildung und damit die Frage des inhaltlichen Zuschnitts, (2) die im Forschungsprozess eingesetzten Referenztheorien und Forschungsmethoden sowie das Kompetenzverständnis, (3) die Bedeutung eines Kompetenzmodells für die Gestaltung des Politikunterrichts (Goll, 2019) sowie (4) die inhaltliche Ausgestaltung eines Kompetenzmodells und hierbei vor allem die Rolle des politischen Wissens (Sander, 2021b).

Hierauf aufbauend und auch unter Beachtung internationaler Forschung zu Kompetenz in der politischen Bildung wurde das Civic Literacy Kompetenzmodell entwickelt. Der Begriff Civic Literacy und die damit verbundene Fokussierung auf politische *und* gesellschaftliche Aspekte stellt eine zumindest teilweise Versöhnung der politikwissenschaftlichen (Detjen, et al., 2012) und der sozialwissenschaftlichen (Autorengruppe Fachdidaktik, 2011) Perspektive in der Kompetenzdebatte der politischen Bildung dar. Zudem schließt der Begriff an den in den Studien der OECD gebräuchlichen Literalitätsbegriff (Gogolin, 2004) und die in den USA vom NAEP sowie die international von der IEA durchgeführten Civic-Studien an (National Assessment Governing Board, 2018; Schulz et al., 2022).

Wie schon beim Modell der Politikkompetenz und im ICCS liegt dem entwickelten Kompetenzmodell ein relativ weiter pädagogisch-psychologischer Kompetenzbegriff zugrunde (siehe auch Marci-Boehncke et al., 2023). Das Civic Literacy Kompetenzmodell enthält einen politischen Wissenstest, der eine kognitive Dimension begründet, sowie drei affektiv-behaviorale Dimensionen, bestehend Motivationen, Einstellungen und Volitionen. Zudem wird für den politischen Wissenstest zwischen den drei Teilfähigkeiten *identifizieren und beschreiben, erklären und analysieren* und *argumentieren und abwägen* sowie den drei Inhaltsbereichen *Strukturen, Prozesse* und *Werte/Inhalte* unterschieden. Die Differenzierung der Teilfähigkeiten erfolgt auf Grundlage des im NAEP durchgeführten Civics Assessments (National Assessment Governing Board, 2018) und die Inhaltsbereiche orientieren sich an der allgemeinen Trias der Politikwissenschaft (Blum & Schubert 2018), die sich auch in den Basiskonzepten des Modells *Konzepte der Politik* von Weißeno et al. (2010) wiederfinden. Innerhalb der Dimension politische Motivation werden die Facetten politisches Interesse und politisches Effektivitätsgefühl unterschieden. Das Effektivitätsgefühl schließt internes und externes politisches Effektivitätsgefühl ein. Für die Dimension der politischen Einstellung werden die Facetten Einstellung zur

Demokratie als stärker politikwissenschaftlich bedeutsam und Einstellung zur pluralen Gesellschaft als eher gesellschaftlich bedeutsam unterschieden. Schließlich beinhaltet die Dimension politische Volition die politische und zivilgesellschaftliche Partizipationsbereitschaft.

Das entwickelte Kompetenzmodell und dessen Teilbereiche dienen als Ausgangspunkt für die systematische Erfassung von Civic Literacy an Schulen der Sekundarstufe I in Deutschland. Eine derartige Verknüpfung von kognitiven und affektiv-behavioralen Teilbereichen kann jedoch über die ursprünglich intendierte Zielgruppe hinaus interessant sein. Beispielsweise können die spezifizierten Teilbereiche auch für Forschungsvorhaben im Primarbereich hilfreich sein (siehe Marci-Boehncke et al., 2023). Insbesondere die Erweiterung des Forschungsfokus von kognitiven Dispositionen auf Handlungsdispositionen in den Bereichen Motivation, Einstellung und Volition kann neue Erkenntnisse liefern (Marci-Boehncke et al., 2023). Darüber hinaus kann das Modell in dieser Form und unter Berücksichtigung der kontextualen Bedingungen auch auf die politische Bildung anderer Länder übertragen werden. Beispielsweise wurde das Modell auf den ungarischen Bildungskontext angewandt und dafür genutzt, Probleme und Herausforderungen der politischen Bildung in Ungarn zu identifizieren (Holle & Ványi, 2022). Als Probleme wurden ein geringer Wissenstand, geringes Interesse, geringes institutionelles Vertrauen, antidemokratische Haltungen, Polarisierung und Intoleranz sowie geringe politische Beteiligung und soziale Aktivität identifiziert (Holle & Ványi, 2022). In dem Beitrag wird verdeutlicht, dass sich diese Probleme den Teilbereichen des Civic Literacy Kompetenzmodells zuordnen lassen (Holle & Ványi, 2022).

Eine empirische Überprüfung des Modells belegte, dass die angenommene Zusammenhangsstruktur von den erhobenen Daten weitgehend gestützt wird. Ausnahmen sind die politische und zivilgesellschaftliche Partizipationsbereitschaft, für die ein gemeinsamer Faktor geeigneter ist sowie das politische Effektivitätsgefühl, für das eine Unterscheidung in internes und externes politisches Effektivitätsgefühl sinnvoll ist. Auch in ausgewählten früheren Studien wurde schon zwischen internem und externem Effektivitätsgefühl unterschieden (Ardèvol-Abreu et al., 2020; Oberle, 2018). Für die Partizipationsbereitschaft muss trotz entsprechender theoretischer Vorüberlegungen (Ekman & Amnå, 2012) konsterniert werden, dass die Daten eine Unterscheidung in die Faktoren *politisch* und *zivilgesellschaftlich* nicht stützen. Ein Grund hierfür könnte sein, dass Schüler*innen der Sekundarstufe I zwar einzelne Partizipationsformen unterscheiden können und bereits Präferenzen entwickelt haben, jedoch noch kein ausgereiftes Verständnis hinsichtlich der Unterscheidung von politischer und zivilgesellschaftlicher Partizipation haben. Ein weiterer Grund könnte sein, dass sich eine Unterscheidung der beiden Faktoren nur für die tatsächliche Partizipation, aber nicht für die intendierte Partizipation zeigt.

Darüber hinaus zeigten die Analysen, dass für das Modell skalare Messinvarianz zwischen der 7. und 10. Klassenstufe besteht, weshalb die Mittelwerte der beiden Klassenstufen miteinander verglichen werden können (siehe auch Robitzsch & Lüdtke, 2023). Ein Vergleich der Mittelwerte zeigte, dass die Mittelwerte aller Konstrukte in der 10. Klassenstufe höher lagen als in der 7. Klassenstufe, was die Annahme der Entwicklungssensitivität des Modells zumindest vorläufig unterstützt. Da die Berechnung der Mittelwertunterschiede jedoch auf Basis zweier unabhängiger querschnittlicher Teilstichproben erfolgt, kann dem Modell eine Entwicklungssensitivität nicht abschließend bescheinigt werden. Statistisch signifikante Mittelwertunterschiede wurden für das politische Wissen sowie die beiden politischen Einstellungen gefunden. Durch die Aggregation von Werten können Mittelwertvergleiche aber dazu führen, dass intraindividuelle Entwicklungsdynamiken aufgrund von gleichzeitig auftretenden Zu- und Abnahmen in der beobachtbaren Merkmalsausprägung maskiert werden (für ein Beispiel anhand von politischem Interesse siehe Russo & Stattin, 2017). Für den Fall, dass längsschnittliche Daten vorliegen, ist es daher sinnvoll, neben Mittelwertunterschieden auch Korrelationskoeffizienten zwischen den Messzeitpunkten zu berechnen.

Schließlich wurde für die Überprüfung der konvergenten und diskriminanten Validität der Konstrukte ein nomologisches Netzwerk identifiziert. Auf Basis des spezifizierten Netzwerks wurden Korrelationen zwischen den Konstrukten des Modells sowie einem theoretisch nicht mit Civic Literacy verbundenen Konstrukt berechnet. Als theoretisch unverbundenes Konstrukt wurde das *Gefühl sozialer Zugehörigkeit* gewählt. Auch hier konnten die Analysen die zuvor angenommenen Zusammenhänge weitgehend bestätigen. Einzig die Korrelation zwischen der Einstellung zur pluralen Gesellschaft und des politischen Interesses war entgegen der Erwartung geringer als die Korrelation zwischen der Einstellung zur pluralen Gesellschaft und dem Gefühl sozialer Zugehörigkeit. Der Unterschied zwischen den beiden Korrelationen ist statistisch jedoch nicht signifikant. Ein Grund könnte sein, dass sich sowohl die Einstellung zur pluralen Gesellschaft als auch das Gefühl sozialer Zugehörigkeit einen stärker sozialen Bezug aufweist, als das politische Interesse.

Prinzipiell weisen die empirischen Ergebnisse darauf hin, dass das Modell sowohl in der 7. als auch in der 10. Klassenstufe zur Erfassung von Civic Literacy eingesetzt werden kann. Außerdem ist anzunehmen, dass das Modell und die operationalisierten Inhalte auch in den angrenzenden Klassenstufen zum Einsatz kommen können.

Eine besondere Herausforderung für die Operationalisierung und Erfassung der verschiedenen Bereiche des Civic Literacy Kompetenzmodells stellt das politische Wissen dar.

Anders als die anderen Bereiche des Kompetenzmodells kann das politische Wissen nicht mittels Selbstberichten erfasst werden (siehe *Beitrag V*). Das Kompetenzmodell sieht stattdessen vor, das politische Wissen mit einem Test bestehend aus Single-Choice Items zu erfassen. Wie in der IRT postuliert, wird angenommen, dass die Schwierigkeit eines Items einerseits von der durchschnittlichen Merkmalsausprägung unter den untersuchten Personen und andererseits von der Schwierigkeit des Iteminhalts abhängt (Moosbrugger et al., 2020). Gerade im Kontext politischen Wissens kann der Expositionsgrad der Iteminhalte (d. h. die Wahrscheinlichkeit, mit der die Schüler*innen mit den im Item enthaltenen Konzepten in Kontakt gekommen sind) bedeutsam für die Itemschwierigkeit sein. Das politische Lernen gilt als sprachintensives Lernen (Drippe, 2018). Das Erlernen und Anwenden von Begriffen der Fachsprache ist ein wichtiger Bestandteil des Lernens im Kontext politischer Bildung (GPJE, 2004). Schüler*innen kommen zudem auch außerhalb der Schule regelmäßig mit politischen Inhalten in Kontakt (z. B. über Nachrichtensendungen oder Diskussionen mit Familienangehörigen und Peers) und lernen so, politische Begriffe anzuwenden und zu verstehen (z. B. Hoskins et al., 2017).

Eine Möglichkeit den Expositionsgrad von in Iteminhalten vorkommenden Konstrukten näherungsweise zu bestimmen, sind durchschnittliche Worthäufigkeiten (Brysbaert et al., 2018). Dabei kann zwischen Worthäufigkeiten in unterschiedlichen Sprachsettings unterschieden werden. Im Kontext von Lernsituationen und konkret auch für das politische Wissen bietet sich eine Unterscheidung in alltägliche und akademische Sprachsettings an. In *Beitrag III* wurde untersucht, inwiefern die Schwierigkeit eines Items mit der durchschnittlichen Häufigkeit der im Item vorkommenden Wörter in alltäglichen und akademischen Sprachsettings zusammenhängt.

Die Ergebnisse zeigen, dass Items mit einer hohen Worthäufigkeit in alltäglichen Sprachsettings leichter waren als Items mit einer niedrigen Worthäufigkeit in alltäglichen Sprachsettings. Das Ergebnis ließ sich jedoch nicht auf Worthäufigkeiten in akademischen Sprachsettings übertragen. Eine hohe Worthäufigkeit in akademischen Sprachsettings hing nicht mit der Schwierigkeit eines Items zusammen. Bei gemeinsamer Betrachtung der beiden Indikatoren der Worthäufigkeit in einem Modell zeigt sich, dass die Worthäufigkeit in alltäglichen Sprachsettings weiter statistisch signifikant negativ mit der Itemschwierigkeit zusammenhing. Die Worthäufigkeit in akademischen Sprachsettings hing in dem Modell aber statistisch signifikant positiv mit der Itemschwierigkeit zusammen. Konkret bedeutet das, dass relative Unterschiede zwischen Worthäufigkeiten in alltäglichen und akademischen Sprachsettings zusätzliche Varianz in der Schwierigkeit von Items aufklärten.

4.1.2 Inwiefern stehen unterschiedliche Merkmale der schulischen politischen Bildung mit den verschiedenen Bereichen von Civic Literacy in Zusammenhang?

Die schulische politische Bildung hat zum Ziel, politische Inhalte zu vermitteln und politische Motivationen, Einstellungen und Volitionen zu unterstützen (Kultusministerkonferenz, 2018). Von einem gelingenden Politikunterricht wird daher erwartet, dass er die genannten Dispositionen fördert und so zur Civic Literacy der Schüler*innen beiträgt. Eine Frage, die sich deshalb stellt, ist, welche Merkmale zeichnen einen qualitativvollen Politikunterricht aus. Während für den Unterricht allgemein und auch für einzelne Fächer, wie zum Beispiel den Mathematikunterricht, bereits konzeptuelle Modelle und empirische Evidenz zur Unterrichtsqualität vorliegen (für eine Übersicht siehe Praetorius & Charalambous, 2018), ist für die politische Bildung bisher vergleichsweise wenig beforscht, was gelingenden Unterricht ausmacht (siehe jedoch Weißeno & Grobshäuser, 2021). Außerdem ist bislang nur wenig darüber bekannt, inwiefern unterschiedliche Unterrichtsqualitätsmerkmale mit verschiedenen Zielen des Politikunterrichts zusammenhängen (siehe jedoch Bayram-Özdemir et al., 2016).

Vor diesem Hintergrund wurden Analysen durchgeführt, mit denen untersucht wurde, inwiefern unterschiedliche domänenspezifische Qualitätsmerkmale des Politikunterrichts mit verschiedenen Teilbereichen des Civic Literacy Kompetenzmodells verknüpft sind. Konkret wurde untersucht, inwiefern die unterschiedlichen Qualitätsmerkmale mit politischer und zivilgesellschaftlicher Partizipationsbereitschaft als mittelbares Ziel politischer Bildung (Oberle & Schwanholz, 2023) zusammenhängen und ob diese Zusammenhänge durch die unmittelbaren Unterrichtsziele politisches Interesse und politisches Wissen mediiert werden. Sowohl das offene Unterrichtsklima als auch die kognitive Aktivierung hingen positiv mit der Partizipationsbereitschaft zusammen. Es wurde jedoch kein Zusammenhang zwischen dem Diskutieren aktueller politischer Themen und der Partizipationsbereitschaft gefunden. Der Zusammenhang zwischen der kognitiven Aktivierung und der Partizipationsbereitschaft wurde durch politisches Interesse und politisches Wissen mediiert. Zudem hing das Diskutieren aktueller politischer Themen positiv mit dem politischen Interesse zusammen. Schließlich hing das offene Unterrichtsklima negativ mit dem politischen Wissen zusammen. Gründe hierfür könnten sein, dass ein sehr offenes Unterrichtsklima die Wissensvermittlung erschwert oder Schüler*innen mit einem stark ausgeprägten politischen Wissen höhere Ansprüche an die Offenheit von Diskussionen im Klassenzimmer stellen.

Die Ergebnisse beziehen sich auf den deutschen Bildungskontext, aber konnten in Teilen im indonesischen Bildungskontext repliziert werden (Lonto & Umbase, 2022). Darüber

hinaus dienen die theoretische Rahmung sowie die Ergebnisse der Studie als Referenz für Studien zur politischen Bildung in China (Liu & Yang, 2023), in Schweden (Stattin & Russo, 2023), in Spanien (Martínez-Cousinou et al., 2022), in Österreich (Graf et al., 2023) und in den USA (Chaney, 2022, Martin, 2022).

Hieraus ergeben sich neue Anknüpfungspunkte. Eine Möglichkeit ist neben den absoluten Merkmalsausprägungen auch relative Unterschiede innerhalb einer Lerngruppen beziehungsweise Klasse zu fokussieren. Die relative Merkmalsausprägung kann Aufschluss über die politische Reputation einer*s Schülerin*s geben, was wiederum die spätere politische Partizipation voraussagt (Stattin & Russo, 2023). Auch in Beitrag II wurde argumentiert, dass die Bedeutung relativer Merkmalsausprägungen dafür verantwortlich sein könnte, dass der Unterschied zwischen der 7. und 10. Klassenstufe für das politische Wissen stärker ausgeprägt ist als für die partizipationsnahen Dispositionen (z. B. Partizipationsbereitschaft, Effektivitätsgefühl). Ein derartiger Befund wäre auch für die Vermittlung von Civic Literacy im Klassenraum und die Unterrichtsqualität von Bedeutung. Ein besonderes Augenmerk müsste auf die Förderung von partizipationsnahen Dispositionen solcher Schüler*innen gelegt werden, die innerhalb ihrer Klasse ein vergleichsweise schwach ausgeprägtes politisches Wissen haben. Zukünftige Studien können hier zusätzliche Evidenz liefern, insbesondere hinsichtlich der Fairness unterschiedlicher Qualitätsmerkmale im Politikunterricht, das heißt der Wirksamkeit von Unterricht für Schüler*innen aus unterschiedlichen Leistungsspektren.

Ein weiterer Anknüpfungspunkt ist die Betrachtung von Emotionen im Kontext politischer Bildung. Auf Basis einer Studie mit $N = 594$ Schüler*innen aus Österreich konnte gezeigt werden, dass im Kontext von politischen Diskussionen empfundene Emotionen mit politischen Kontroll- (z. B. politisches Effektivitätsgefühl) und Werturteilen (z. B. politisches Interesse) korrelieren (Graf et al., 2023). Zudem wurden statistisch signifikante Korrelationen zwischen einzelnen Emotionen und dem Engagement beziehungsweise dem politischen Wissen der untersuchten Schüler*innen gefunden. Statistisch signifikante Korrelationen wurden für die Emotionen Freude, Hoffnung und Stolz (positiv) sowie für Scham, Hoffnungslosigkeit und Langeweile (negativ) festgestellt. Weniger eindeutig waren die Ergebnisse für die Emotionen Wut und Angst. Zukünftige Forschung könnte einen Beitrag leisten, das Verhältnis zwischen unterschiedlichen Qualitätsmerkmalen des Politikunterrichts und im Unterricht empfundene Emotionen aufzuklären (Graf et al., 2023).

Insgesamt verdeutlichen die in Beitrag IV vorgestellten Forschungsergebnisse sowie auf den Beitrag bezugnehmende Studien die Wichtigkeit eines zieldifferenten Einsatzes von Qualitätsmerkmalen und Methoden im Politikunterricht. Theoretische Modelle sowie empirische

Evidenz zur Unterrichtsqualität im Kontext schulischer politischer Bildung sind im Vergleich zu einigen anderen Fachdidaktiken (Praetorius & Gräsel, 2021b) bislang jedoch weniger stark ausgebildet. Eine simple Übertragung der generischen Tiefenstrukturen qualitativollen Unterrichts scheint den besonderen Gegebenheiten des Politikunterrichts, wie zum Beispiel der notwendigerweise diskursiven Auseinandersetzung mit Themen, der Bedeutung von Meinung und Meinungsbildung sowie einem etwaigen Bezug zu aktuellen Nachrichtenthemen, nicht gerecht zu werden. Zukünftige Forschung sollte daher einen Beitrag zur Theoriebildung für die Tiefenstrukturen qualitativollen Politikunterrichts leisten.

4.1.3 Wie hängen subjektives und objektives politisches Wissen zusammen??

Der Zusammenhang zwischen selbsteingeschätzter und tatsächlicher Fähigkeit ist von besonderer Bedeutung für den Lern- und Bildungsprozess von Individuen und Gegenstand unterschiedlicher Forschungsbereiche (z. B. Andrade, 2019; Dunlosky & Rawson, 2012; Yan & Carless, 2022). Insbesondere im Kontext der politischen Bildung können fehlerhafte Selbsteinschätzungen Konsequenzen haben, die über den Lern- und Bildungsprozess hinausgehen (z. B. radikale Einstellungen, van Prooijen & Krouwel, 2020).

Mittels Selbsteinschätzungen kann die Wahrnehmung der eigenen Performanz in einem Test oder die eigene Fähigkeit ausgedrückt werden. Der Dunning-Kruger Effekt beschreibt ein Befundmuster systematischer Verzerrung zwischen Selbsteinschätzungen und tatsächlicher Fähigkeit beziehungsweise Testleistung (Kruger & Dunning, 1999), bei dem sich leistungsschwache Individuen weniger genau einschätzen als leistungsstarke Individuen und leistungsschwache Individuen eher zu einer Überschätzung und leistungsstarke Individuen eher zu einer Unterschätzung neigen.

Die in Beitrag V durchgeführten Analysen zeigen, dass der Dunning-Kruger Effekt auch in Bezug auf das politische Wissen bei Schüler*innen in den Klassenstufen 7 und 10 gefunden wurde. Das heißt, dass Schüler*innen, die in dem entwickelten politischen Wissenstest schlecht abschnitten, ihr politisches Wissen weniger genau einschätzten als Schüler*innen, die in dem Wissenstest gute Ergebnisse erzielten. Zudem neigten Schüler*innen mit einem schlechten Testergebnis eher zu einer Überschätzung und Schüler*innen mit einem guten Testergebnis eher zu einer Unterschätzung. Die von Kritiker*innen des Dunning-Kruger Effekts empfohlene statistische Überprüfung mittels Varianzanalyse (Gignac & Zajenkowski, 2020) bestätigt den visuellen Eindruck, dass sich die Differenz zwischen selbsteingeschätztem politischen Wissen und der Leistung im Wissenstest statistisch signifikant zwischen den vier Leistungsquartilen der jeweiligen Klassenstufe unterschied. Diese Ergebnisse verdeutlichen, dass die Erfassung

politischen Wissens mittels Selbstberichten nicht nur ungenaue, sondern auch systematisch verzerrte Werte erzeugt und eine objektive Erfassung, wie im entwickelten Civic Literacy Kompetenzmodell vorgesehen, notwendig ist.

Außerdem wurde gefunden, dass der Dunning-Kruger Effekt in den beiden untersuchten Klassenstufen 7 und 10 in sehr ähnlicher Weise auftritt. Auch hier bestätigte eine Varianzanalyse den visuellen Eindruck, dass es keine statistisch signifikanten Unterschiede zwischen den Klassenstufen gibt. Die Beständigkeit des Effekts ist insbesondere vor dem Hintergrund der besonderen Entwicklungsdynamik in der Zeit zwischen der 7. und 10. Klassenstufe interessant. Im Allgemeinen ist die frühe Adoleszenz eine Zeit, die durch einen dynamischen Reifungsprozess gekennzeichnet ist (Dahl et al., 2018) und die sich durch neurobiologische Veränderungen und Konsequenzen für das kognitive Leistungsvermögen (Laube et al., 2020), die metakognitiven Fähigkeiten (Weil et al., 2013) sowie den Lern- und Bildungserfolg (Goering et al., 2023) auszeichnet. Auch im Kontext der politischen Sozialisation ist die Adoleszenz eine wichtige Entwicklungsphase. Der originäre Politikunterricht setzt für gewöhnlich in der frühen Adoleszenz ein und Schüler*innen beginnen erstmals damit, sich mit politischen Themen außerhalb ihrer Familie auseinanderzusetzen (Hoskins et al., 2017). Der Einfluss des Elternhauses geht in diesem Alter zurück und Jugendliche entwickeln stärker unabhängige Einstellungen (Grob, 2009). Zudem ist die Modifizierbarkeit politischer Orientierungen und Verhaltensweisen vergleichsweise hoch in der Adoleszenz (siehe „Impressionable Years“-Hypothese in der Entwicklungs- und Persönlichkeitsforschung, Visser & Krosnick, 1998). Die Entwicklung erfolgt dabei nach dem Prinzip einer sukzessiven Kristallisation, in der sich das Wissensniveau, politische Orientierungen und Verhaltensweisen stabilisieren. Im weiteren Lebensverlauf nimmt die Plastizität der Einstellungen und Orientierungen ab (Baumert et al. 2016).

Dass der Dunning-Kruger Effekt dennoch in sehr ähnlicher Form in der 7. und 10. Klassenstufe auftritt, könnte ein Indiz dafür sein, dass es sich beim Dunning-Kruger Effekt um ein statistisches Artefakt handelt und systematische Unterschiede in den metakognitiven Fähigkeiten von Schüler*innen nicht für den Effekt verantwortlich sind (Gignac & Zajenkowski, 2023; Hiller, 2023). Zwar kann eine abschließende Klärung der Bedeutung metakognitiver Fähigkeiten für den Dunning-Kruger Effekt zum jetzigen Zeitpunkt nicht erfolgen, jedoch demonstrieren die Ergebnisse, dass der Effekt auch im Kontext politischer Bildung existiert und die Rolle metakognitiver Fähigkeiten für den Effekt weiterhin ungeklärt ist.

Zudem wurde untersucht, wie das wahrgenommene Lehrkräfteverhalten mit der Einschätzung der Schüler*innen zusammenhängt. Ähnlich zu den Befunden in Beitrag IV wird

deutlich, dass die wahrgenommene Unterrichtsqualität im Kontext politischer Bildung für verschiedene Unterrichtsziele eine unterschiedliche Bedeutsamkeit hat und ein zieldifferenter Einsatz unterschiedlicher Techniken, Methoden und Verhaltensweisen notwendig ist. Konkret zeigen die Ergebnisse, dass ein als motivierend wahrgenommener Unterricht mit einer genaueren Selbsteinschätzung, jedoch auch mit einer höheren Wahrscheinlichkeit der Überschätzung einherging. Außerdem wurde gefunden, dass ein als kognitiv aktivierend wahrgenommener Unterricht mit einer weniger genauen Selbsteinschätzung zusammenhing. Die Ergebnisse zur Motivationsqualität stimmten mit den zuvor formulierten Hypothesen überein. Die negative Beziehung zwischen kognitiver Aktivierung und Urteilsgenauigkeit entspricht jedoch nicht der zuvor formulierten Hypothese. Ein Grund hierfür könnte sein, dass Schüler*innen, die ihren Unterricht als besonders kognitiv aktivierend wahrnehmen, höhere Erwartungen an sich selbst und ihren bisherigen Lernerfolg stellen und ihre bisherige Lernleistung daher überschätzen.

Durch die Verwendung eines innovativen Zwei-Kohorten-Designs, die Verknüpfung der klassischen visuellen Untersuchung des Dunning-Kruger Effekts mit verschiedenen statistischen Methoden sowie die Berücksichtigung des Lernsettings wird eine neue Perspektive auf den Dunning-Kruger Effekt eröffnet.

Zukünftige Forschung kann mithilfe des entwickelten Wissenstests und aufgrund der Möglichkeit, diesen in längsschnittlichen Untersuchungen einzusetzen, an diese Studie anknüpfen. Eine Möglichkeit wäre eine konkretere Erfassung der Selbsteinschätzungen. Zum Beispiel könnte die Selbsteinschätzung in Form der Anzahl der als richtig eingeschätzten Antworten erfasst werden. Außerdem wäre eine Einschätzung der eigenen Leistung im Vergleich zu den Peers (z. B. Mitschüler*innen innerhalb der eigenen Klasse) eine Möglichkeit, das selbsteingeschätzte politische Wissen zu erfassen.

Zudem kann die Fähigkeit der Selbsteinschätzung mit weiteren Facetten des Civic Literacy Kompetenzmodells verknüpft werden. Zum Beispiel kann untersucht werden, inwiefern die Überschätzung des eigenen politischen Wissens mit den übrigen Facetten des Kompetenzmodells zusammenhängt. Auch kann untersucht werden, welche Faktoren neben dem Lehrkräfteverhalten Überschätzung und Unterschätzung bedingen.

Insgesamt verdeutlichen die Ergebnisse, dass die Untersuchung des Zusammenhangs zwischen subjektivem und objektivem politischen Wissen wichtige Implikationen für die Beforschung des Dunning-Kruger Effekts sowie die pädagogische Praxis in der politischen Bildung birgt.

4.2 Stärken und Limitation

Im Folgenden werden die übergreifenden Stärken und Limitationen dieser Arbeit erläutert, die bei der Betrachtung der Ergebnisse und deren Interpretation beachtet werden sollten. Hierbei wird jeweils zwischen Stärken und Limitation, die das Studiendesign oder das methodische Vorgehen betreffen, unterschieden.

4.2.1 Stärken

Zu den das Studiendesign betreffenden Stärken gehören die ganzheitliche Betrachtung von Civic Literacy, die den klassischen Forschungsprozess überspannende Instrumentenentwicklung, -erprobung und -überprüfung und die Verknüpfung unterschiedlicher Wissenschaftsdisziplinen.

Das dieser Arbeit zugrundeliegende Verständnis von Civic Literacy geht von einem multidimensionalen Konstrukt aus und unterstreicht die Notwendigkeit, im Kontext der Beforschung von Civic Literacy verschiedene Dimensionen und Facetten zu berücksichtigen. Wenn es darum geht, festzustellen, inwiefern junge Menschen in der Lage sind, sich im politischen und gesellschaftlichen System zurechtzufinden und sich in demokratischer Weise daran zu beteiligen, reicht es nicht aus, sich ausschließlich das politische Wissen oder das politische Interesse anzusehen (Carretero et al., 2016; Luengo-Kanacri et al., 2023). Das entwickelte Modell stellt somit einen Orientierungsrahmen für Kompetenzziele im Kontext politischer Bildung dar – insbesondere im Rahmen der schulischen Bildung in der Sekundarstufe I. Vorteile des Modells gegenüber anderen Modellen sind, dass alle Modellbereiche vollständig operationalisiert und empirisch erprobt sind, alle Dimensionen und Facetten des Modells auf einer jeweils gleichen modelltheoretischen Ebene liegen (d. h., dass zumindest theoretisch keine einseitigen Abhängigkeiten bestehen) und dass alle Instrumente des Modells prinzipiell für einen Einsatz in längsschnittlichen Untersuchungen geeignet sind.

Eine weitere Stärke ist, dass die Arbeit einen ganzheitlichen quantitativen Forschungsprozess wiedergibt, bei dem aufbauend auf theoretischen Vorüberlegungen und empirischen Erkenntnissen ein Civic Literacy Kompetenzmodell konzeptualisiert, operationalisiert und zur Datenerhebung eingesetzt wurde (*Beitrag I, II und III*; Reinders, 2022). Anschließend wurden die erhobenen Daten ausgewertet, interpretiert und zur weiteren Verbesserung des Kompetenzmodells genutzt (*Beitrag II*). Schließlich wurden die entwickelten Instrumente zur Beforschung inhaltlicher Fragestellungen eingesetzt (*Beitrag IV und V*).

Außerdem zeichnet sich die Arbeit durch einen interdisziplinären Forschungscharakter aus. Neben der Pädagogischen Psychologie, sowie der Empirischen Bildungsforschung finden

auch Theorien und Methoden der Politikdidaktik und Politikwissenschaft Berücksichtigung. Eine besondere Rolle spielt hierbei die Zuwendung zu und Integration von empirischen Konstrukten der Pädagogischen und Kognitionspsychologie (Weißeno, 2012). Die interdisziplinäre Herangehensweise dieser Arbeit erweitert das Verständnis von Civic Literacy und eröffnet neue Möglichkeiten für die Beforschung von politischer Bildung bei jungen Menschen.

Bezogen auf das methodische Vorgehen sind der Einsatz verschiedener Analysemethoden, der Einsatz neuartiger und reliabler Messinstrumente sowie die Nutzung aktueller Daten von insgesamt $N = 1.297$ Schüler*innen als Stärken hervorzuheben.

Zur Untersuchung der Forschungsfragen in den verschiedenen Beiträgen dieser Arbeit wurden unterschiedliche und jeweils für das Ziel des Beitrags angemessene Analysemethoden verwendet. Für die Skalierung der Daten des politischen Wissenstests wurde ein unidimensional multigroup 1-parameter Modell gerechnet, mit dem es möglich ist, die Fähigkeit der Schüler*innen unterschiedlicher Klassenstufen sowie die Schwierigkeit der einzelnen Items zu bestimmen. Hierfür wurden zuvor Ankeritems bestimmt, die von Schüler*innen beider Klassenstufen bearbeitet wurden. Zudem wurden Prozessdaten genutzt, um *rapid-guessing-behavior* zu identifizieren und entsprechende Antworten im Vorfeld der Skalierung auszuschließen. Für die Beantwortung der Forschungsfragen wurden unter anderem *confirmatory factor analyses*, eine *ordinary least squares regression*, *structural equation models* sowie *one-way* und *two-way ANOVAs* gerechnet.

Im Kontext der Modellentwicklung wurden teilweise neue Messinstrumente entwickelt und empirisch erprobt. Die neu entwickelten Messinstrumente weisen zufriedenstellende bis gute Reliabilitäten in beiden Klassenstufen auf. Auch die Diskriminanz und Konvergenz der im Kompetenzmodell eingesetzten Messinstrument wird durch die empirischen Daten gestützt.

Schließlich wurden im Kontext dieser Arbeit Daten von insgesamt $N = 1.297$ Schüler*innen erhoben. Im Rahmen der ersten Studie wurden insgesamt $N = 250$ Schüler*innen (*Beitrag III* und *IV*) und im Rahmen der zweiten Studie $N = 1.047$ Schüler*innen (*Beitrag II* und *V*) befragt. Die Datensätze beinhalten demographische Angaben, die Ergebnisse des politischen Wissenstests, Angaben zu den übrigen Bereichen des entwickelten Kompetenzmodells sowie weiteren politischen Einstellungen und Orientierungen, Informationen zur Wahrnehmung des Politikunterrichts sowie zu außerschulischen Lerngelegenheiten und dem familiären Umfeld. Darüber hinaus wurden auch das Leseverständnis und die Lesegeschwindigkeit der Schüler*innen erfasst und ein Matrizenest eingesetzt. Zudem ist die Möglichkeit der längsschnittlichen Weiterverfolgung der Schüler*innen gegeben. Bezogen auf das Geschlecht und den Migrationshintergrund sind die erhobenen Daten der zweiten Studie mit der tatsächlichen

Verteilung in NRW vergleichbar (IT.NRW, 2022). Zudem wurden die Daten für die Analysen in Beitrag II und V mit Blick auf die Schulformzugehörigkeit der Schüler*innen gewichtet.

4.2.2 Limitationen

Neben den Stärken dieser Studie gibt es jedoch auch einige Limitationen. Bezogen auf das Studiendesign sind die querschnittliche Datenstruktur, die durch geringen Leistungsdruck gekennzeichnete Testumgebung sowie der pandemische Kontext, in dem die Daten erhoben wurden, zu nennen.

Prinzipiell können querschnittliche und längsschnittliche Daten zu Beantwortung unterschiedlicher Forschungsfragen genutzt werden. Nicht immer ist das Sammeln längsschnittlicher Daten notwendig. Zwar können längsschnittliche Daten zusätzliche Evidenz hinsichtlich der prognostischen Bedingungsstruktur unterschiedlicher Variablen liefern, allerdings kann auch hieraus nicht automatisch Kausalität abgeleitet werden. Insbesondere im Kontext von Beitrag II (Forschungsfrage 3 zur Entwicklungssensitivität) und Beitrag II (Forschungsfrage 2 zum Unterschied des Dunning-Kruger Effekts in Klassenstufe 7 und 10) wären längsschnittliche Daten jedoch eine Bereicherung gewesen.

Anders als in high-stakes Testungen hatten die Ergebnisse keine unmittelbaren Konsequenzen für die Schüler*innen. Die Datenerhebung kann daher als low-stakes Testung verstanden werden (Köller et al., 2019). Im Vorfeld der Datenerfassung wurde den Schüler*innen gegenüber transparent gemacht, dass weder die Studienteilnahme noch die gegebenen Antworten Gegenstand einer Benotung oder anderweitigen Leistungsbewertung durch eine Lehrkraft sind und dass alle Antworten pseudonymisiert werden. Das hat zur Folge, dass die Ergebnisse nicht ohne Weiteres auf reale beziehungsweise high-stakes Testsituationen übertragen werden können. Im Gegensatz zu high-stakes Testsituationen sind low-stakes Testsituationen durch eine geringere Leistungsmotivation und geringere Testängstlichkeit gekennzeichnet (Knehta, 2017). Aufgrund der Vergleichbarkeit der Testumgebung sind die Ergebnisse jedoch in besonderem Maße relevant für LSAs in diesem Bereich.

Die Daten der dieser Arbeit zugrundeliegenden Beiträge wurden im Winter 2020/2021 (Studie I; *Beitrag III* und *IV*) und im Winter 2021/22 (Studie II, *Beitrag II* und *V*) erhoben und fallen daher in einen Zeitraum, in dem der Bildungsbereich in Deutschland in besonderem Maße durch die COVID-19-Pandemie beeinträchtigt war (z. B. Schulschließungen, Wechselunterricht, Abstandsregelung). Es wurden jedoch alle Daten in Präsenz im Klassenraum erhoben. Vor, während und im Anschluss an die Datenerhebungen wurden stets die zu der Zeit geltenden Pandemievorschriften und Hygieneregeln eingehalten. Es ist anzunehmen, dass insbesondere die

zeitweisen Schulschließungen den Lernfortschritt der Schüler*innen beeinflusst haben (z. B. Betthäuser et al., 2023). Zudem scheinen sozio-ökonomisch benachteiligte Schüler*innen vergleichsweise stärker von den Pandemiebedingungen betroffen gewesen zu sein (Betthäuser et al., 2023).

Auch mit Blick auf das methodische Vorgehen gibt es einige Limitationen. Hierzu gehören der Umgang mit der hierarchischen Struktur der Daten sowie die Repräsentativität der beiden Stichproben.

Die befragten Schüler*innen lassen sich jeweils einer Klasse und einer Schule zuordnen. Um die hierarchische Struktur der Daten zu berücksichtigen sind spezielle Analyseverfahren notwendig. In den Beitragsanalysen wurden stets robuste Standardfehler berechnet, indem die Klassenzugehörigkeit als Clustervariable spezifiziert wurde. Es muss weiter davon ausgegangen werden, dass die Zugehörigkeit zu einer Klasse oder einer Schule Einfluss auf Merkmalsausprägungen der Schüler*innen hat. Mithilfe von Mehrebenen-Analysen kann geprüft werden, inwiefern die Gruppenzugehörigkeit im Zusammenhang mit der Merkmalsausprägung einer*s Schüler*in steht. Insbesondere können mittels Mehrebenen-Analysen die anteilige Varianzaufklärung unterschiedlicher Ebenen sowie die Bedeutung von Prädiktoren unterschiedlicher Ebenen bestimmt werden. Voraussetzung hierfür sind jedoch ausreichend große Stichproben auf den höheren Ebenen (z. B. Anzahl Klassen oder Schulen). Im Kontext politischer Bildung könnten Mehrebenen-Analysen Aufschluss darüber geben, inwiefern die Zugehörigkeit zu einer Klasse oder einer Schule mit den unterschiedlichen Bereichen des Civic Literacy Kompetenzmodells und deren Niveauveränderungen zusammenhängt.

Die für die Analysen genutzten Daten gehören zu einer Stichprobe, die repräsentativ für eine Grundgesamtheit stehen soll. Die in dieser Studie interessierende Grundgesamtheit umfasst Schüler*innen der Sekundarstufe I in den Klassenstufen 7 und 10 an allgemeinbildenden Schulen, insbesondere in NRW. Idealerweise erfolgt die Auswahl der Schüler*innen für die Stichprobe randomisiert aus der Grundgesamtheit. Aus forschungspraktischen Gründen wurden jedoch nicht individuelle Schüler*innen, sondern Schulen für die Stichprobe ausgewählt. Da die ausgewählten Schulen zu keiner Teilnahme an der Studie verpflichtet waren, ist von einer Beeinträchtigung der Repräsentativität aufgrund von Selbstselektion auf Schulebene auszugehen. Hinzu kommt, dass auch die Schüler*innen nicht zu einer Teilnahme verpflichtet waren und daher auch schülerseitig Selbstselektion eine Rolle gespielt haben könnte. In einem Teil der vorgenommenen Analysen konnte die Repräsentativität der Stichprobe mithilfe der Gewichtung der Schulformzugehörigkeit verbessert werden. Denkbar wäre zudem eine

Gewichtung entlang sozio-ökonomischer Merkmale. Zudem könnte die Repräsentativität durch größere Stichproben weiter verbessert werden.

4.3 Implikationen für Forschung und Praxis

4.3.1 Implikationen für die Forschung

Die vorliegende Dissertation leistet einen wichtigen Beitrag zur systematischen Beforschung von Civic Literacy, indem ein Civic Literacy Kompetenzmodell konzeptualisiert, operationalisiert und empirisch erprobt wurde (*Beitrag I, II und III*). Auf Basis empirischer Daten wurden die zuvor getätigten Modellannahmen überprüft und notwendige Anpassungen vorgenommen (*Beitrag II*). Das so entwickelte Kompetenzmodell kann zur systematischen Beforschung von politischer Bildung und Civic Literacy im Kontext der Sekundarstufe I eingesetzt werden. Insbesondere die Operationalisierung der Modellinhalte stellt einen wichtigen Schritt dar, um den an die schulische politische Bildung formulierten Wirkungsanspruch (Kultusministerkonferenz, 2018) empirisch überprüfen zu können. Mithilfe der Entwicklung altersangemessener und psychometrisch überprüfter Messinstrumente können Prädiktoren, Niveau und Entwicklung von Civic Literacy empirisch überprüft werden.

Anschließend an Beitrag II und den darin enthaltenen Analysen ist eine flächendeckende und systematische Erfassung von Civic Literacy zur Bestimmung von Niveau und Entwicklung in Deutschland ein wichtiges Desiderat. Diesem Desiderat wird durch den zukünftigen Einsatz der Modellinstrumente im Nationalen Bildungspanel (NEPS) nachgekommen. Die Instrumente werden im Kontext des NEPS insbesondere in den Klassenstufen 7 und 10 eingesetzt und bilden somit einen Zeitraum ab, der durch dynamische Veränderungen in der Civic Literacy junger Menschen gekennzeichnet ist (Baumert et al. 2016 sowie *Beitrag II* in dieser Arbeit). Durch die im NEPS geplante große Stichprobe und umfängliche Datengrundlage mit Rückmeldungen unterschiedlicher Befragungsgruppen werden neue Analysen ermöglicht, die zur Bearbeitung wichtiger Forschungsdesiderate beitragen können.

In Anschluss an die Ergebnisse in Beitrag III ist die Untersuchung des Zusammenhangs zwischen Lesekompetenz und Civic Literacy ein Desiderat für zukünftige Forschung. Das Erlernen und Beherrschen eines Fachwortschatzes ist insbesondere im Kontext der politischen Bildung als vergleichsweise sprachintensives Unterrichtsfach wichtiger Bestandteil des Lernprozesses (Drippe, 2018; GPJE, 2004). Von besonderem Interesse wäre daher, inwiefern Lesekompetenz den Erwerb von Civic Literacy begünstigt und welche Bereiche des Civic Literacy Kompetenzmodells auch unabhängig von Lesekompetenz vermittelt werden können.

Außerdem könnten die Entwicklung und der Einsatz eines politischen Wortschatztests Aufschluss darüber geben, inwiefern das Erlernen politischer Begriffe tatsächlich prädiktiv für die unterschiedlichen Bereiche des Civic Literacy Kompetenzmodells ist.

Ein weiteres aus dieser Arbeit abzuleitendes Forschungsdesiderat ist die Weiterentwicklung der in Beitrag IV unternommenen Systematisierung von Unterrichtsqualität im Kontext politischer Bildung. Während eine Anpassung des Modells der generischen Unterrichtsqualität in einigen Schulfächern – insbesondere in naturwissenschaftlichen Fächern (Praetorius et al. 2020) – bereits weit fortgeschritten ist (Praetorius & Gräsel, 2021b), ergibt sich für theoretisch abgeleitete und empirisch überprüfte Modelle der Unterrichtsqualität in der politischen Bildung bisher eine Angebotslücke (siehe jedoch Weißeno & Grobshäuser, 2021). Die Analysen in Beitrag IV könnten entsprechend um weitere Unterrichtsqualitätsmerkmale erweitert und deren Einfluss auf die unterschiedlichen Bereiche des Civic Literacy Kompetenzmodells untersucht werden.

Das Analyseverfahren in Beitrag V stellt eine wichtige Erweiterung bisheriger Forschungsdesigns zur Ergründung des Dunning-Kruger Effekts dar. Der Einsatz eines Zwei-Kohorten-Designs kann auch in zukünftigen Studien zum Dunning-Kruger Effekt einen Beitrag zur Verständniserweiterung liefern. Ein mögliches Desiderat für zukünftige Forschungsprojekte zum Dunning-Kruger Effekt im Kontext politischer Bildung ist eine veränderte Operationalisierung der Selbsteinschätzung. Denkbar wären beispielsweise relative Selbsteinschätzungen, bei der die Testteilnehmer*innen ihre selbsteingeschätzte Testleistung in Form von Rankperzentilen ausdrücken. Vergangene Studien zeigten, dass der Effekt stärker in Erscheinung tritt, wenn relative Einschätzungen (Rankperzentile) anstelle von absoluten Einschätzungen (Anzahl richtiger Antworten) genutzt werden (McIntosh et al., 2019). Zudem könnte ein Vergleich der vor und nach dem Test gegebenen Selbsteinschätzungen (*predictions* vs. *postdictions*) interessant sein (Miller & Geraci, 2011). Ein weiteres Desiderat besteht in der näheren Betrachtung von Testteilnehmer*innen mit einer inakkuraten Selbsteinschätzung. Wichtige Fragen wären beispielsweise, welche individuellen Faktoren für inakkurate Selbsteinschätzungen verantwortlich sind und welche Konsequenzen inakkurate Selbsteinschätzungen für das politische Lernen und die Bereiche des Civic Literacy Kompetenzmodells haben.

Über die Einzelbeiträge hinaus bestehen weitere Desiderate für zukünftige Forschungsprojekte. Eine besondere Rolle kommt dabei der längsschnittlichen Erfassung von Civic Literacy zu. Bislang liegt nur zu einzelnen Abschnitten und einzelnen Bereichen Evidenz zur

Entwicklungsdynamik von Civic Literacy im Jugendalter vor (Kuger & Giller, 2020). Der Einsatz des Civic Literacy Kompetenzmodells in einer Panelstudie wie dem NEPS wird daher wichtige Erkenntnisse hinsichtlich der Entwicklungsdynamik von Civic Literacy im Jugendalter sowie den Bedingungsbeziehungen des Modells liefern. Ein weiteres Desiderat ist die Beforschung des familiären Hintergrunds und dessen Bedeutung für den Erwerb von Civic Literacy. Bisherige Forschung zeigte, dass Ungleichheiten im politischen Engagement von Schüler*innen insbesondere auf Unterschiede in der familiären Sozialisation zurückzuführen seien (z. B. Deimel et al., 2020; Neundorf & Smets, 2017). Vergleichsweise wenig Evidenz liegt bisher zu der Frage vor, inwiefern sich unterschiedliche Merkmale des familiären Hintergrunds auf die unterschiedlichen Bereiche des entwickelten Civic Literacy Kompetenzmodells auswirken. Im Bereich der politischen Sozialisation werden elterliche Einflüsse und die Übertragung unterschiedlicher politischer Orientierungen als intergenerationale Transmission thematisiert (Grob, 2009). Aufgrund der prinzipiellen Übertragbarkeit der Modellbereiche eignet sich das Civic Literacy Kompetenzmodell auch für die Untersuchung etwaiger Transmissionseffekte. Auch kann durch die gemeinsame Betrachtung der schulischen und familiären Sozialisation eine Beurteilung der relativen Wichtigkeit der beiden Instanzen für die verschiedenen Bereiche des Kompetenzmodells erfolgen (Neundorf et al., 2016).

4.3.2 Implikationen für die Praxis

Neben den Implikationen für die Forschung bergen die Befunde dieser Arbeit auch Implikationen für die pädagogische Praxis und insbesondere die politische Bildung in der Sekundarstufe I.

Maßnahmen zur Förderung der demokratischen Kultur in Deutschland haben in Form verschiedener Bundesprogramme längst Eingang in die deutsche Bildungspolitik gefunden (z. B. „kompetent. für demokratie“, „Toleranz fördern – Demokratie stärken“, „Demokratie leben!“, „Land.schafft.Demokratie“). Insbesondere Schulen haben in Deutschland den Auftrag, junge Menschen auf ein Leben in der Demokratie vorzubereiten und wichtige demokratische und politische Fähigkeiten zu vermitteln (Kultusministerkonferenz, 2018). Tagtäglich findet an deutschen Schulen gelungene Demokratieförderung und politische Bildung sowohl im dafür vorgesehenen Fachunterricht als auch als Querschnittsaufgabe statt. Allerdings gibt es auch in Deutschland weiter Optimierungsbedarf. Die Ergebnisse der in Beitrag II vorgenommenen Analysen zeigen, dass sich das durchschnittliche Niveau des politischen Effektivitätsgefühls, des politischen Interesses und der Partizipationsbereitschaft kaum zwischen Schüler*innen der

7. und 10. Klassenstufe unterscheidet. Zudem gaben nur etwa 61 Prozent der befragten Schüler*innen an, in Zukunft an politischen Wahlen teilnehmen zu wollen (K7: 58 %; K10: 65 %). Abgesehen davon offenbarten die Ergebnisse der ICCS 2016 ein im europäischen Vergleich unterdurchschnittliches politisches Wissen der Schüler*innen in NRW (Abs & Hahn-Laudenberg, 2017).

Das entwickelte Civic Literacy Kompetenzmodell kann als Orientierungsrahmen einen wichtigen Beitrag für die Planung und Zielsetzung politischer Bildung an Schulen leisten. Das Modell verdeutlicht, dass sich gelingender Politikunterricht nicht alleine auf Wissensvermittlung verstehen kann, sondern weitere Dispositionen in den Bereichen Motivation, Einstellung und Volition berücksichtigt werden sollten (Luengo-Kanarci et al., 2023). Die Befunde in Beitrag II legen nahe, dass sich die Merkmalsausprägungen der Schüler*innen in den einzelnen Modellbereichen in unterschiedlicher Weise zwischen der 7. und 10. Klassenstufe ändern. Insbesondere das politische Wissen sowie die Zustimmung zur Demokratie und zur pluralen Gesellschaft scheinen in dieser Zeit zuzunehmen (siehe auch Baumert et al., 2016; Janmaat & Hoskins, 2022).

Beitrag IV zeigt zudem, dass das durch die Schüler*innen wahrgenommene Lehrkräfteverhalten in unterschiedlicher Weise mit den verschiedenen Modellbereichen zusammenhängt. Während die kognitive Aktivierung positiv mit dem politischen Interesse, dem politischen Wissen und der Partizipationsbereitschaft der Schüler*innen zusammenhing, wurde für das Diskutieren aktueller politischer Themen ein positiver Zusammenhang mit dem politischen Interesse, aber nicht mit dem politischen Wissen oder der Partizipationsbereitschaft gefunden. Für das offene Unterrichtsklima bestand zwar ein positiver Zusammenhang mit der Partizipationsbereitschaft, aber ein negativer Zusammenhang mit dem politischen Wissen. Diese Ergebnisse unterstreichen die Notwendigkeit eines zieldifferenten Einsatzes von Unterrichtsmethoden in der politischen Bildung (Maurissen, 2020; Sanjaya et al., 2022; Tegelbecker, 2023).

Auch die Befunde in Beitrag V bestärken die unterschiedliche Bedeutung verschiedener Unterrichtsmethoden. Ein als motivierend wahrgenommenes Lehrkräfteverhalten hing mit genaueren Einschätzungen des eigenen politischen Wissens zusammen. Allerdings neigten Schüler*innen, die ihre Lehrkraft als sehr motiviert wahrnehmen, häufiger zu einer Überschätzung ihres politischen Wissens. Den nicht hypothesenkonformen Befund, dass ein als kognitiv aktivierend empfundener Unterricht mit weniger genauen Einschätzungen zusammenhing, gilt es in weiteren Studien zu überprüfen und den verantwortlichen Mechanismus zu ergründen, ehe hieraus Implikationen für die Praxis abgeleitet werden sollten.

4.4 Fazit

Die Einführung von Bildungsstandards in der BRD zu Beginn der 2000er Jahre hat auch in der Politikdidaktik zu einer Diskussion über die Notwendigkeit national verbindlicher Bildungsstandards und die inhaltliche Ausgestaltung eines Kompetenzmodells für die politische Bildung geführt (Goll, 2019). Das in dieser Dissertation entwickelte und empirisch überprüfte Civic Literacy Kompetenzmodell schließt an die Diskussion sowie an nationale und internationale Forschungsarbeiten im Kontext schulischer politischer Bildung an (z. B. Abs & Hanhn-Laudenberg, 2017; Detjen et al., 2012; National Assessment Governing Board; Schulz et al., 2022). In Beitrag I wurde Civic Literacy aufbauend auf theoretischen Vorüberlegungen und empirischen Untersuchungen definiert und in die Dimensionen politisches Wissen, politische Motivationen, politische Einstellungen und politische Volitionen unterteilt. In den Beiträgen II und III wurden theoretische Annahmen des entwickelten Kompetenzmodells und des politischen Wissenstests empirisch überprüft und wo immer nötig entsprechende Änderungen vorgenommen. In Beitrag IV wurde untersucht, inwiefern unterschiedliche Unterrichtsmethoden mit verschiedenen Bereichen des Kompetenzmodells zusammenhängen. Schließlich wurde in Beitrag V die Beziehung zwischen objektiv gemessenem und selbsteingeschätztem politischen Wissen sowie dessen Zusammenhang mit motivierendem und kognitiv aktivierendem Politikunterricht analysiert. Das in den Beiträgen I, II und III entwickelte und empirisch überprüfte Kompetenzmodell kann zukünftig eingesetzt werden, um die Civic Literacy von Schüler*innen der Sekundarstufe I systematisch zu erfassen. Der Einsatz des Modells wird neben der Bestimmung des Niveaus und der Entwicklung von Civic Literacy auch die Identifikation von Faktoren, die die Vermittlung von Civic Literacy positiv oder negativ beeinflussen, ermöglichen. Die Beiträge IV und V zeigen beispielhaft auf, in welchen Forschungskontexten die entwickelten Instrumente potentiell zum Einsatz kommen können. Die Ergebnisse der Beiträge IV und V verdeutlichen zudem, dass ein zieldifferenter Einsatz von Unterrichtsmethoden in der politischen Bildung nötig ist, da nicht jede Unterrichtsmethode für jedes Ziel des entwickelten Kompetenzmodells geeignet ist. Die Ergebnisse können als Hinweis auf die besondere Bedeutung von Lehrkräften und den Einsatz unterschiedlicher Unterrichtsmethoden für gelingende politische Bildung gesehen werden. Durch den Einsatz des entwickelten Civic Literacy Kompetenzmodells in zukünftigen Forschungsprojekten und der weiteren Untersuchung politischer Bildung und dessen Gelingensbedingungen kann ein wichtiger Beitrag für die Beforschung und Praxis politischer Bildung geleistet werden.

Literaturverzeichnis II

- Abs, H. J. & Hahn-Laudenberg, K. (Hrsg.). (2017). *Das politische Mindset von 14-Jährigen: Ergebnisse der International Civic and Citizenship Education Study 2016*. Waxmann.
- Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education, 4*. <https://www.frontiersin.org/articles/10.3389/feduc.2019.00087>
- Ardèvol-Abreu, A., Gil de Zuniga, H. & Gámez, E. (2020). The influence of conspiracy beliefs on conventional and unconventional forms of political participation: The mediating role of political efficacy. *British Journal of Social Psychology, 59*(2), 549–569. <https://doi.org/10.1111/bjso.12366>
- Autorengruppe Fachdidaktik (Hrsg.). (2011). *Politik und Bildung: Bd. 64. Konzepte der politischen Bildung: Eine Streitschrift*. Wochenschau Verlag.
- Baumert, J., Becker, M., Cortina, K. S., Köller, O., Kropf, M. & Maaz, K. (2016). Die Entwicklung des politischen Interesses und des Selbstkonzepts der politischen Kompetenz vom Jugend- bis in das Erwachsenenalter. In C. Grunert, A. Schippling & N. Pfaff (Hrsg.), *Kritische Bildungsforschung: Standortbestimmungen und Gegenstandsfelder* (S. 323–352). Barbara Budrich.
- Bayram-Özdemir, S., Stattin, H. & Özdemir, M. (2016). Youth's initiations of civic and political discussions in class: Do youth's perceptions of teachers' behaviors matter and why? *Journal of Youth and Adolescence, 45*(11), 2233–2245. <https://doi.org/10.1007/s10964-016-0525-z>
- Bethhäuser, B. A., Bach-Mortensen, A. M. & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour, 7*(3), 375–385. <https://doi.org/10.1038/s41562-022-01506-4>
- Blum, S. & Schubert, K. (2018). *Politikfeldanalyse: Eine Einführung* (3. Aufl.). Springer. <https://doi.org/10.1007/978-3-658-17758-4>
- Brysbaert, M., Mander, P. & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science, 27*(1), 45–50. <https://doi.org/10.1177/0963721417727521>
- Carretero, M., Haste, H. & Bermudez, A. (2016). Civic education. In L. Corno & E. M. Anderman (Hrsg.), *Handbook of educational psychology* (S. 309–322). Routledge.
- Chaney, L. E. (2022). Civic Education Initiatives in Central Texas: A Descriptive Case Study.

- Dahl, R. E., Allen, N. B., Wilbrecht, L. & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, 554(7693), 441–450. <https://doi.org/10.1038/nature25770>
- Deimel, D., Hoskins, B. & Abs, H. J. (2020). How do schools affect inequalities in political participation: compensation of social disadvantage or provision of differential access? *Educational Psychology*, 40(2), 146–166. <https://doi.org/10.1080/01443410.2019.1645305>
- Detjen, J., Massing, P., Richter, D. & Weißeno, G. (2012). *Politikkompetenz – ein Modell*. Springer.
- Dewey, J. (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan. <https://books.google.de/books?id=OlqwHwAACAAJ>
- Drippe, M. (2018). *Der Fachwortschatz in Schulbüchern für die 9. und 10. Klasse des Politikunterrichts an Gesamtschulen*. Cives! Scholl of Civic Education.
- Dunlosky, J. & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Ekman, J. & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3), 283–300. <https://doi.org/10.2478/s13374-012-0024-1>
- Gignac, G. E. & Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data. *Intelligence*, 80, 101449. <https://doi.org/10.1016/j.intell.2020.101449>
- Goering, M., Albright, M. G. & Mrug, S. (2023). The Effects of Pubertal Timing on Academic Performance in Adolescence and Career Success in Adulthood: Evidence from a 16-year Longitudinal Study. *Journal of Youth and Adolescence*, 52(9), 1769–1787. <https://doi.org/10.1007/s10964-023-01814-6>
- Gogolin, I. (2004). Zum Problem der Entwicklung von „Literalität“ durch die Schule. *Zeitschrift für Erziehungswissenschaft*, 7(3), 101–111.
- Goll, T. (2019). Die Rezeption des Modells der Politikkompetenz in der Politikdidaktik. In G. Weißeno (Hrsg.), *Politik lernen: Studien und theoretische Ansätze* (S. 25–39). Springer. https://doi.org/10.1007/978-3-658-27896-0_3
- GPJE. (2004). *Anforderungen an Nationale Bildungsstandards für den Fachunterricht in der Politischen Bildung an Schulen: Ein Entwurf*. Wochenschau Verlag. <http://gpje.de/wp-content/uploads/2017/01/Bildungsstandards-1.pdf>

- Graf, E., Goetz, T., Bieleke, M. & Murano, D. (2023). Feeling Politics at High School: Antecedents and Effects of Emotions in Civic Education. *Political Psychology*. Vorab-Onlinepublikation. <https://doi.org/10.1111/pops.12907>
- Grob, U. (2009). Die Entwicklung politischer Orientierung vom Jugend- ins Erwachsenenalter: - Ist die Jugend eine spezifisch sensible Phase in der politischen Sozialisation? In H. Fend, U. Grob & F. Berger (Hrsg.), *Lebensverläufe, Lebensbewältigung, Lebensglück: Ergebnisse der Life-Studie* (S. 305–326). Springer .
- Hiller, A. (2023). Comment on Gignac and Zajenkowski, “The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data”. *Intelligence*, 97, 101732. <https://doi.org/10.1016/j.intell.2023.101732>
- Holle, A. & Ványi, É. (2022). Helyzetkép a hazai iskolai Állampolgári Nevelésről: Kontextus, kudarcok, kutatási irányok és kihívások. *Politikatudományi Szemle*, 16(3). <https://doi.org/10.30718/POLTUD.HU.2022.3.7>
- Hoskins, B., Janmaat, J. G. & Melis, G. (2017). Tackling inequalities in political socialisation: A systematic analysis of access to and mitigation effects of learning citizenship at school. *Social science research*, 68, 88–101. <https://doi.org/10.1016/j.ssresearch.2017.09.001>
- IT.NRW. (2022). *Schüler/-innen*) mit Zuwanderungsgeschichte an allgemeinbildenden und beruflichen Schulen nach Verwaltungsbezirk und Schulform*. https://www.it.nrw/sites/default/files/itnrw_presse/326_22.pdf
- Janmaat, J. G. & Hoskins, B. (2022). The Changing Impact of Family Background on Political Engagement During Adolescence and Early Adulthood. *Social Forces*, 101(1), 227–251. <https://doi.org/10.1093/sf/soab112>
- Knekta, E. (2017). Are all Pupils Equally Motivated to do Their Best on all Tests? Differences in Reported Test-Taking Motivation within and between Tests with Different Stakes. *Scandinavian Journal of Educational Research*, 61(1), 95–111. <https://doi.org/10.1080/00313831.2015.1119723>
- Köller, O., Hasselhorn, M., Hesse, F. W., Maaz, K., Schrader, J., Solga, H., Spieß, C. K. & Zimmer, K. (2019). *Das Bildungswesen in Deutschland: Bestand und Potenziale*. utb.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037//0022-3514.77.6.1121>
- Kultusministerkonferenz. (2018). *Demokratie als Ziel, Gegenstand und Praxis historisch-politischer Bildung und Erziehung in der Schule*. KMK.

- https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschlu-esse/2009/2009_03_06-Staerkung_Demokratieerziehung.pdf
- Laube, C., van den Bos, W. & Fandakova, Y. (2020). The relationship between pubertal hormones and brain plasticity: Implications for cognitive training in adolescence. *Developmental Cognitive Neuroscience*, 42, 100753. <https://doi.org/10.1016/j.dcn.2020.100753>
- Liu, L. & Yang, L. (2023). Investigating the use of fuzzy logical data for evaluating the teaching quality of college teachers. *Journal of Intelligent & Fuzzy Systems*, 44(6), 10459–10475.
- Lonto, A. L. & Umbase, R. S. (2022). The Integration of Political Interest in Transmitting the Teaching Quality Management of Civics Education in Indonesia. *Eurasian Journal of Educational Research*, 99, 233–248.
- Luengo-Kanacri, B. P., Jiménez-Moya, G., Miranda, D. & Marinovic, A. (2023). When civic knowledge matters but is not enough: The role of classroom climate and citizenship self-efficacy on different facets of civic engagement. *Citizenship Teaching & Learning*, 18(1), 119–140. https://doi.org/10.1386/ctl_00113_1
- Marci-Boehncke, G., Rath, M. O. & Tkotzyk, R. (2023). Media, Language and their Impact on the Development of Young Children's Political Awareness. In M. Carmo (Hrsg.), *Education Applications & Developments VIII: Advances in Education and Educational Trends Series* (S. 419–434). inScience Press.
- Martin, T.-R. (2022). *Black Americans at a Crossroads in American Politics: Electoral Participation, Protest, and Policy Change*. University of Missouri-Saint Louis.
- Martínez-Cousinou, G., Camus-García, E. & Álvarez-Sotomayor, A. (2022). Juventud universitaria e interés por la política: Análisis de un estudio piloto. 1885-589X. <https://doi.org/10.46661/revintpensampolit.7504>
- Maurissen, L. (2020). Political efficacy and interest as mediators of expected political participation among Belgian adolescents. *Applied Developmental Science*, 24(4), 339–353. <https://doi.org/10.1080/10888691.2018.1507744>
- McIntosh, R. D., Fowler, E. A., Lyu, T. & Della Sala, S. (2019). Wise up: Clarifying the role of metacognition in the Dunning-Kruger effect. *Journal of experimental psychology. General*, 148(11), 1882–1897. <https://doi.org/10.1037/xge0000579>
- Miller, T. M. & Geraci, L. (2011). Unskilled but aware: reinterpreting overconfidence in low-performing students. *Journal of experimental psychology. Learning, memory, and cognition*, 37(2), 502–506. <https://doi.org/10.1037/a0021802>

- Moosbrugger, H., Gäde, J. C., Schermelleh-Engel, K. & Rauch, W. (2020). Klassische Testtheorie (KTT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 275–304). Springer. https://doi.org/10.1007/978-3-662-61532-4_13
- National Assessment Governing Board. (2018). *Civics Framework: for the 2018 National Assessment of Educational Progress*. Washington, DC. National Assessment Governing Board, U.S. Department of Education. <https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/civics/2018-civics-framework.pdf>
- Neundorf, A., Niemi, R. G. & Smets, K. (2016). The compensation effect of civic education on political engagement: How civics classes make up for missing parental socialization. *Political Behavior*, 38(4), 921–949. <https://doi.org/10.1007/s11109-016-9341-0>
- Neundorf, A. & Smets, K. (2017). Political Socialization and the Making of Citizens. In Oxford Handbooks Editorial Board (Hrsg.), *Oxford Handbook Topics in Politics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199935307.013.98>
- Oberle, M. (2018). Politisches Effektivitätsgefühl von Schüler/-innen. In S. Manzel & M. Oberle (Hrsg.), *Kompetenzorientierung* (S. 85–97). Springer. https://doi.org/10.1007/978-3-658-16889-6_8
- Oberle, M. & Schwanholz, J. (2023). *Demokratie ohne Wahlbeteiligung? Potenziale politischer Bildung an Schulen besser nutzen!* Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/bueros/nrw/19927.pdf>
- Praetorius, A.-K. & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM*, 50(3), 535–553. <https://doi.org/10.1007/s11858-018-0946-0>
- Praetorius, A.-K. & Gräsel, C. (Hrsg.). (2021). Unterrichtsqualität – fachspezifisch oder generisch? Ein Update zu Perspektiven aus der Fachdidaktik und der allgemeinen Lehr-Lernforschung [Sonderheft]. *Unterrichtswissenschaft*, 49(2). Springer.
- Reinders, H. (2022). Überblick Forschungsmethoden. In H. Reinders, D. Bergs-Winkels, A. Prochnow & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 153–159). Springer. https://doi.org/10.1007/978-3-658-27277-7_9
- Robitzsch, A. & Lüdtke, O. (2023). Why Full, Partial, or Approximate Measurement Invariance Are Not a Prerequisite for Meaningful and Valid Group Comparisons. *Structural equation modeling: a multidisciplinary journal*, 1–12. <https://doi.org/10.1080/10705511.2023.2191292>

- Russo, S. & Stattin, H. (2017). Stability and change in youths' political interest. *Social indicators research*, 132(2), 643–658. <https://doi.org/10.1007/s11205-016-1302-9>
- Sander, W. (2009). Wie standardisierbar ist Bildung? Chancen und Probleme von Bildungsstandards in Deutschland. In C. Buschkühle, L. Duncker & V. Oswalt (Hrsg.), *Bildung zwischen Standardisierung und Heterogenität: ein interdisziplinärer Diskurs* (S. 11–33). Springer.
- Sander, W. (2021). Kompetenzorientierung als Forschungs- und Konfliktfeld der Didaktik der politischen Bildung. In W. Sander & K. Pohl (Hrsg.), *Handbuch politische Bildung* (S. 122–132). Wochenschau Verlag.
- Sanjaya, D. B., Suartama, I. K. & Suastika, I. N. (2022). The Effect of the Conflict Resolution Learning Model and Portfolio Assessment on the Students' Learning Outcomes of Civic Education. *International journal of instruction*, 15(1), 473–488.
- Schulz, W., Fraillon, J., Losito, B., Agrusti, G., Ainley, J., Damiani, V. & Friedman, T. (2022). *International Civic and Citizenship Education Study 2022: Assessment Framework*. Springer.
- Stattin, H. & Russo, S. (2023). The Political Reputation of Students in Upper Secondary School: Consequences for Their Collective Political Activities Four Years Later. *Political Psychology*, 44(4), 693–708. <https://doi.org/10.1111/pops.12879>
- Teegelbeckers, J. Y., Nieuwelink, H. & Oostdam, R. J. (2023). School-based teaching for democracy: A systematic review of teaching methods in quantitative intervention studies. *Educational Research Review*, 100511. <https://doi.org/10.1016/j.edurev.2023.100511>
- Weißeno, G. (Hrsg.). (2012). *Schriftenreihe / Bundeszentrale für Politische Bildung: Bd. 1191. Politisch handeln: Modelle, Möglichkeiten, Kompetenzen*. Bundeszentrale für politische Bildung.
- Weißeno, G., Detjen, J., Juchler, I., Massing, P. & Richter, D. (2010). *Konzepte der Politik - ein Kompetenzmodell. Schriftenreihe: Bd. 1016*. Bundeszentrale für politische Bildung. https://www.pedocs.de/volltexte/2016/12009/pdf/Weisseno_et_al_2010_Konzepte_der_Politik_.pdf
- Weißeno, G. & Grobshäuser, N. (2021). Unterrichtsqualität des Politikunterrichts in der gymnasialen Oberstufe. In C. Deichmann & M. Partetzke (Hrsg.), *Demokratie im Stresstest: Reaktionen von Politikdidaktik und politischer Bildung* (S. 249–264). Springer. https://doi.org/10.1007/978-3-658-33077-4_14

Yan, Z. & Carless, D. (2022). Self-assessment is about more than self: the enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, 47(7), 1116–1128. <https://doi.org/10.1080/02602938.2021.2001431>