

Political judgement competency among upper secondary-school pupils

Weißeno, Simon; Weißeno, Georg

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Weißeno, S., & Weißeno, G. (2021). Political judgement competency among upper secondary-school pupils. *Citizenship Teaching & Learning*, 16(1), 7-27. https://doi.org/10.1386/ctl_00044_1

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-SA Lizenz (Namensnennung-Nicht-kommerziell-Weitergabe unter gleichen Bedingungen) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-SA Licence (Attribution-NonCommercial-ShareAlike). For more information see: <https://creativecommons.org/licenses/by-nc-sa/4.0>

Citizenship Teaching & Learning
Volume 16 Number 1

© 2021 Intellect Ltd Article. English language. https://doi.org/10.1386/ctl_00035_1

Received 6 January 2020; Accepted 10 June 2020

SIMON WEISSENO
Alfons Kern Schule

GEORG WEISSENO
University of Education Karlsruhe

Political judgement competency among upper secondary-school pupils

ABSTRACT

Within the discipline of political didactics, differing views exist on political judgement as the goal or content of classroom teaching. In this study, political judgement is understood as a competency. It requires situational deliberation and decision-making, but must also take into account political values. For this study, 401 upper secondary pupils in years 11–12 were presented with a 45-minute judgement test. The pupils were asked to adopt a personal position on a specific matter in an essay, using argumentative deliberation. As in the lens model of judgement, they were required to adduce and evaluate different aspects under conditions of uncertainty, since no definitive information was available. In the test, five levels were used to assess the 'complexity' of a judgement. All previous tasks had to be solved before a higher level of complexity could be reached. On the basis of the assumed interdependencies of levels (tasks) in a testlet, the test was scaled using a testlet model from item response theory. All the testlets show significantly higher variances than the test as a whole. The test was able to produce a good total variance. The analyses of construct validity by means of fluid intelligence and subject knowledge conform to expectations. The test evaluated with the testlet model indicates that this can be assumed to be a multilevel process.

KEYWORDS

political judgement
political competence
political knowledge
fluid intelligence
decision-making
socio-scientific issues
critical thinking
testlet model

In political didactics, political judgement is generally seen as the goal or content of classroom teaching. There are, however, differing views as to what actually constitutes 'good' (i.e. sophisticated) political judgement.

A number of political didactics specialists associate 'good' judgement with a political attitude that is relevant to the topic and has been independently formulated by learners. They understand political judgement as 'attitude formation' linked with a call to political action (e.g. Waghid 2005; Brunold 2017). Others associate 'good' judgement with the selection of facts from lessons or from political reality, a selection justified both normatively and by intrinsic logic, in line with the philosophical concept of rationality (e.g. Juchler 2005; Kolstø 2008).

Still others see the need for operationalizable and non-normative criteria for 'good' judgement, in order to make it empirically measurable and transparently assessable at school (e.g. Weißeno forthcoming). In accordance with this, political judgement is regarded as a competency. Judgement competency is decision-making about specific political situations or issues, after deliberation and justification based on conceptual political knowledge (Manzel and Weisseno 2017). According to this view, judgements involve situational deliberation and decision-making, but must take into account political values. The emphasis is on criteria for judgement which relate to defined school-based domain-specific knowledge. The empirical value of this definition of judgement competency is still largely unproven.

This article focuses on the psychometric quality of a test designed to measure judgement competency at upper secondary level. The article investigates whether the measuring instrument used allows stable and valid measurements of knowledge. The first step will be to introduce concepts of judgement from political didactics, other domain-specific didactics and psychology. The next step will be to explain the understanding on which this study is based, using the model of political judgement competency and the measurements of judgement competency. The research questions, measuring instruments, sample and methods of analysis will then be presented. The subsequent sections will report and discuss the quality of the test and the results.

THEORETICAL BACKGROUND

Reflections within political didactics

In German political didactics, pupils' capacity to judge is a long-standing subject of reflection. Authors have expressed differing ideas about what a political judgement should be. Lingelbach (1970) proposed a model for the classification of political disputes, based on interest, concepts of value and order, the ethics of responsibility and commitment. He is concerned with attitudes and actions, and with a very small number of aspects chosen on a domain-specific basis.

Grosser (1987: 166ff) formulates only content-related aspects of political judgement, with three degrees of difficulty: (1) Sufficient life experience is available, and the experts are essentially in agreement. (2) The issues are complex, but the experts are still essentially in agreement, even if they argue with different interests and priorities. (3) The issues are highly complex, and the experts are not in agreement. Grosser focuses clearly on the content of politics lessons, and highlights the degree of complexity of the judgement. This approach, however, lacks generalizable specifications and criteria for the different degrees of difficulty. Complexity remains a vague criterion.

Massing (1997: 125f.) attempts to combine complexity and content-related criteria in an easy-to-use judgement 'grid', on which the content of politics lessons can be located. This judgement grid, featuring the categories of efficiency and legitimacy, is therefore often used in classroom practice and teacher training, to allow issues to be judged with political rationality. There are three categories to take into account for a rationally justified decision. (1) An issue may be contemplated from the point of view of the political actors, the addressees or the democratic system (point of view category). The point of view, in each case, refers to the position with which a political decision is justified; (2) the evaluative category 'efficiency of political action' is assessed through means-end rationality; and (3) the evaluative category 'legitimacy' stands for e.g. social sustainability. In the process of judgement formation, pupils are expected to reach an individual judgement, emphasizing what is reasonable and generalizable about the points of view. The grid is also intended to allow reflection on the judgements of others.

This proposal with the heuristic of a grid shows how difficult it is to define the quality of a political judgement by means of conceptual analysis. This is because the selected perspective and the subjective appraisal of the (content-neutral) efficiency of a political measure do not have to relate to quality, but may also correspond to preference. The assessment of a measure's legitimacy may have a subjective, prescriptive component, and does not necessarily contain any substance. It is, furthermore, a value-dependent decision. Nor are there any operationalizable indications as to how the quality of rationality is to be measured. A heuristic alone does not make it possible to assess judgements with generalizable criteria. Breit (1997: 132) therefore points out that there is as yet no theory behind the categories and questions on judgement formation.

Another position is to define the capacity for judgement in abstract, socio-logical terms. Proponents of this position formulate new heuristics with even more aspects.

I define it as the ability to adopt a reasoned position within the realm of possibility of political thought, with the help of political measures of value, a) with reference to the specific case, and b) in a fundamentally identity-related way, and to understand the positions adopted by those with different ideas.

(Petrik 2012: 32)

The vague term 'identity-related' is associated with subjective views, beliefs and interests. The proposal gives recommendations for personal identity formation, without offering any specific, operationalizable details.

The outcome of all this is that intersubjectively differing definitions and evaluations of a political judgement are probably the norm among political didactics specialists and teachers. The authors have not given any thought to the operationalizability of the aspects, which are situated on different levels. The models do not yet allow any theory-based empirical research.

For the school sector, the specifications are set out in the standardized examination requirements for the Germany university entrance exam for politics lessons (*Einheitliche Prüfungsanforderungen in der Abiturprüfung für Politik/Sozialkunde*, KMK 2005). The requirements set out by the education administration state what politics lessons at upper secondary-level schools in Germany are to have achieved by the time students finish school. Here political judgement is cited as an ability which must be fostered. In everyday school

life, pupils' judgements are graded. Here judgements contain domain-specific knowledge. They are found at the highest level of the requirements, level three. Pupils are expected to acquire 'systematic and structured knowledge', which enables them to 'form reasoned, independent and thoughtful judgements based on domain-specific knowledge' (KMK 2005: 4). For this, 'criteria or categories of judgement' are necessary (KMK 2005: 5). What remain unclear in the *Einheitliche Prüfungsanforderungen* are the concrete substantive criteria for the diagnosis and subsequent grading of a political judgement. They offer only abstract statements and structural descriptions, which are not operationalized (Goll 2008).

This review of the descriptions of the capacity for judgement which have so far been proposed within political didactics has shown that all such proposals allow for the normativity of judgements (Weisseno 2016). The idea of seeking levels of complexity already exists. The previous approaches emphasize the content of a political judgement to varying degrees.

Reflections within subject-specific didactics and psychology

In international research, efforts have been made to prove the theoretical and at the same time empirical worth of the concepts of judgement. There are studies here which refer to the capacity for judgement mainly in terms of 'valuation' and 'decision-making'. The term 'decision-making' corresponds to the process model of valuation competency, understood as individual decision-making based on knowledge and values (Knittel 2013). In 'decision-making', a distinction is made between individual stores of general knowledge and values and a body of subject-related knowledge which can be built up in subject-specific lessons (Osborne et al. 2004).

In the field of epistemological theories, the most comprehensive schema for 'socio-scientific issues' has been elaborated by King and Kitchener (2002), with their 'reflective judgement model'. Socio-scientific issues, i.e. topics of public controversy such as genetic engineering, nuclear power, or climate protection, often have a scientific background. Solving these 'ill-structured' problems requires the integration of multiperspectival approaches, and the ability to argue and judge on the basis of subject knowledge. The reflective judgement model tries to find answers to the question of how the different developmental stages of children and adolescents can be supported so that they can make independent and rationally justified judgements and decisions in subsequent stages of development.

King and Kitchener assume a multistage structure of subjective concepts of knowledge. Here arguments on the lowest stage of pre-reflective thinking are simple. They can be based on personal perceptions of reality, on authorities such as God, or on absolute truths. Arguments in the middle stages include the factor of uncertainties with regard to absolute truths, but decisions here are often based on personal feelings or general norms. Only in the highest stages of reflective thinking are arguments context-based, plausibility and coherence are applied to lines of argument, and personal value preferences are reflected on rationality. But pupils can draw on ideas and knowledge developed informally through their own life world experiences, and their ethical values (cf. Osborne et al. 2004). This does not yet make it clear, however, that it is political values and facts which should be at the forefront in the stages of political judgement, not personal preference.

Judgement is closely related to critical thinking. Here the focus is on the extent to which epistemological beliefs influence the process of thought and argument, and reflexive judgement. According to the empirical taxonomy of critical thinking developed by Dick (1991), this includes not only the identification and analysis of arguments, but also scientific, analytical inference and logic. When external influences are taken into account, then values, authorities and emotional language are also part of argument. It should be noted that this understanding of judgement includes expertise and contextualization. Comparable modellings of judgement competency are found in history didactics (Trautwein et al. 2017) and German language didactics (Frederking et al. 2011). These are, however, not so readily applicable to political contexts.

In the German research context, judgement has been introduced and to some extent empirically examined in other subjects, as part of the new focus on competence and the formulation of educational standards (cf. Hostenbach et al. 2011; Riemeier et al. 2012). In chemistry didactics, the ESNaS competency model (Evaluation of Standards in the Natural Sciences) divides valuation competency into the sub-areas of 'valuation criteria', 'options for action' and 'reflection', with each area being further differentiated into five levels of increasing complexity (Hostenbach et al. 2011). Here decision-making is subdivided into decision-making situation, decision-making process and decision-making rules. The external influencing factors include personal values systems and social norms (Ajzen and Fishbein 1969: 134). Factors cited as sources of difficulty are increasing complexity and the kind of cognitive processes involved. Like ideas from political didactics, this model often contains an individual moral dimension and thus a normative notion of judgement.

A model of ethical valuation competency from biology brings together eight dimensions of competence. According to Bögeholz et al. (2004), this is not about moral judgement on issues, but about rational justification during an argumentative exchange. The basis of the model is therefore the capacity for argument, which also comprises domain-specific knowledge. However, stores of general and domain-specific knowledge are mixed and not analytically separated from one another. Furthermore, judgement in the natural science models is more about making subjectively influenced normative decisions on particular questions. Political judgement, in contrast, is about engaging with political values on the basis of domain-specific (political) knowledge, and weighing up alternative decision-making options. Political judgements cannot be justified with definitive information.

The models from the natural sciences have as their background the psychological process model of decision-making, with the phases of pre-selection (search for information, need for action), selection (valuation process, separation of facts and personal values) and post-selection (verification of feasibility) (Betsch and Haberstroh 2005). In decision-making theory, a distinction is made between non-compensatory and compensatory decision rules (cf. Jungermann et al. 2005). Of interest for political didactics is the compensatory decision rule: according to this rule, options for action relating to political controversies which score badly or worse than other options on one or more attributes can compensate for this disadvantage if they score well on other attributes (Pfister et al. 2017: 98). They offer a good reflection of a qualitative process of deliberation in domain-specific political argument and judgement. In contrast, this is not possible in the case of non-compensatory rules.

On the one hand, then, the models from the natural sciences indicate that political controversies are very similar to socio-scientific issues, and,

with the capacity for argument, also show a proximity to critical thinking. The aspects mentioned are undoubtedly important for the examination of political judgements in political didactics. Cognitive processes in argument, increasing complexity (levels or stages), personal value preferences, and the idea of decision-making are relevant for political judgements. On the other hand, the models do not yet take into account the specific conditions of political action and valuation in the presence of alternative solutions, and therefore the necessary domain-specific frame of reference.

THE MODEL OF POLITICAL JUDGEMENT COMPETENCY

Politics teaching requires a theory of the capacity for judgement which can be operationalized and therefore tested. This is the only way to ensure that it is anchored in classroom reality. What is needed is a definition of the standards, positions and contents that are to be used for judgement. The modelling of judgements must describe the cognitive demands placed on pupils as they learn and apply concepts and arguments from political science (Weissenro 2015: 162). Political judgement is therefore modelled as a competency in the following discussion (Weissenro 2017a: 12). This approach makes it possible to give a theory-based description of classroom reality.

Since the educational reform of 2003 and the development of educational standards in Germany, the capacity for political judgement has been firmly anchored in the philosophy of the subject. The competence model quickly produced by the relevant professional body (the German Society for Civic Education Didactics and Civic Youth and Adult Education) accorded a prominent position to the capacity for judgement. It calls for fact-based and value-based analysis of the issues, and reflective judgements (GPJE 2004: 13), and it stipulates that the degree of complexity in pupils' reasoning should be considered when it comes to assessment, and that value judgements should aim to be valid for everyone (GPJE 2004: 15). This document formulates normative demands for domain-specific judgements without any theoretical references from the international or psychological debate, and without providing any operationalizable criteria.

The first theoretical derivation of the capacity for political judgement as a psychological concept of competence (Klieme et al. 2008) for politics teaching came with the model of political competence (Detjen et al. 2012). It takes into account both the subject-specific and normative aspects of a judgement. Normative judgements require a valuation or the adoption of a position, while factual judgements are defined as descriptive statements about issues. The distinction between description and valuation constitutes an operationalizable distinction (cf. Weissenro 2017a). However, the five levels of judgement in this model (declaratory judgements, expansion judgements, value judgements, decision judgements and shaping judgements) proved unsuitable in the pilot testing for this study. The structure of the levels and the presumed sequence of the judgement process were not borne out empirically. For the normative element, this model names political values such as peace and justice from the model of subject knowledge. For the description of facts or issues, subject knowledge is implicitly adduced via the linguistic elements, which are thought to be operationalizable.

It should be noted that this model uses only subject-related standards for the assessment of judgements in the classroom (Detjen et al. 2012: 61). The other individual opinions of the pupils are used to attain the rates of learning,

if and when they express technically correct judgements on the current class-room topic. Thus a domain-specific argument and judgement should draw on subject knowledge (Weisseno et al. 2010). On the one hand this contains subject-related concepts such as democracy, state and party, and on the other hand political values such as freedom, equality, justice, peace and sustainability.

The model of political judgement competency on which this study is based (Manzel and Weisseno 2017) is an extrapolation of the model of political competence (Detjen et al. 2012). In this model, political judgement in school is technically correct reasoning, deliberation and decision-making with regard to political facts and issues (Manzel and Weisseno 2017). The model seeks to operationalize the levels of judgement using levels of complexity. Judgement competency, defined in this way, is context-specific and thus content-related, i.e. the relevant terms and concepts from political science must be applied situationally in a judgement.

In a judgement, statements on political reality are evaluated with political standards (Manzel and Weisseno 2017: 72). The pupils make individual decisions on concrete, reality-related issues. They can refer to several facts, or just a few. These must be weighed up against each other, and a decision must be made. In the problem-solving situation at school (e.g. in the completion of a task), conflicting assessments of facts lead to a complex context of justification (cf. Gronostay 2014: 37 ff.). A psychological foundation can be seen, as in the lens model (Brunswik 1955). In this model, a judgement requires different aspects to be considered and weighted under conditions of uncertainty, since definitive information is not available. Argument-based decision-making is therefore the essence of political judgement.

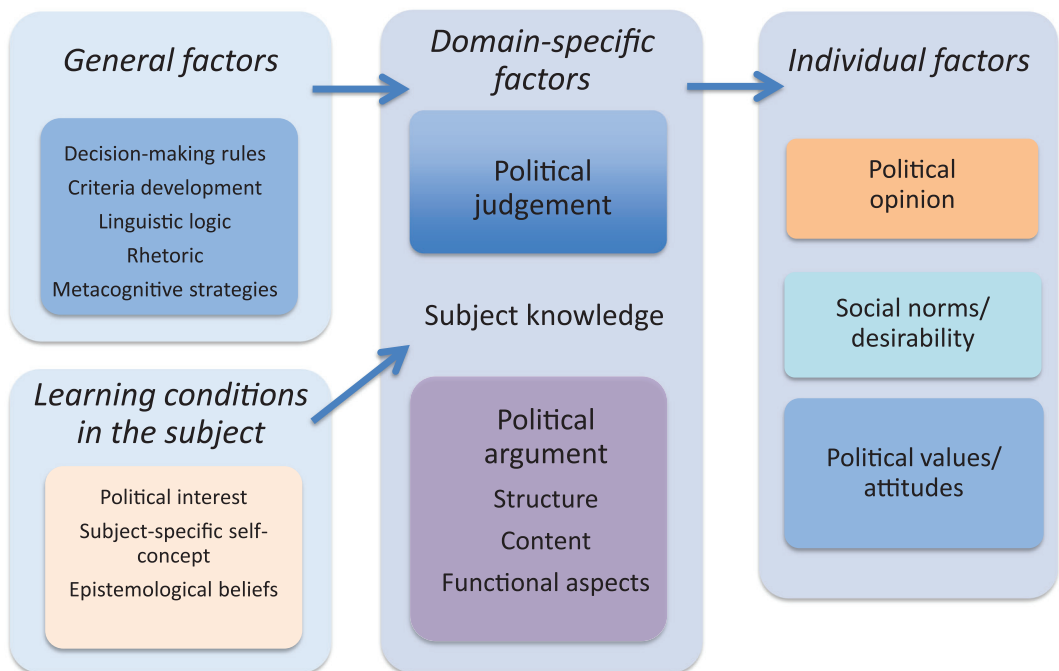


Figure 1: Model of the capacity for political judgement (Manzel and Weisseno 2017: 71).

The model of judgement competency uses the ‘complexity’ of a judgement to describe judgement processes. Complexity can be ascertained by the number of (counter-)arguments/facts and of different political facts used for valuation, and by the process of deliberation and weighting which takes place here (cf. Osborne et al. 2004). Graduations of judgement become possible (Table 1). The five levels or stages allow the graduation of competence to be measured. The quality of a judgement is enhanced by a multitude of deliberation processes in the course of the argument. Here political judgement is based on dependent cognitive processes which build on one another.

As for the natural sciences, the task complexity for political decision-making situations is defined by the quantity of information that must be processed to complete the task, and by the connections between the units of information (Hostenbach et al. 2011: 280). It is the interconnections pupils make within their answers which make it possible to describe the level. Here individual norms and statements of political opinion (everyday knowledge) are not included in the assessment. The levels relate to the standards for politics teaching formulated in the model. Level 5, however, includes broader subject-related facts going beyond classroom teaching.

Besides the graduation of judgement, another aspect to be clarified is the significance of the subject knowledge which is processed in the judgements. Domain-specific content knowledge encompasses 30 subject concepts constituting political knowledge. These concepts are selected on the basis of the model of subject knowledge that should be taught at school (Weisseno et al. 2010: 12). The concepts are democracy, European actors, European integration, freedom, peace, justice, separation of powers, equality, human rights, interest groups, international relations, conflict, legitimacy, power, mass media, market, human dignity, sustainability, public sphere, public goods, opposition, parliament, parties, constitutional state, government, representation, security, welfare state, state and elections. Additional terms are subsumed under these concepts. These concepts and around 220 related technical terms are not presented as isolated units of knowledge to be memorized, but rather as points in a network of knowledge.

The domain-specific concepts make it possible to form political judgements and lines of argument with reference to specific addressees. The pupils can formulate them with the political terms and concepts mentioned. Thus

| Level | |
|--------------|---|
| 0 | Not answered, answered without reference to the question, no decision-making/argument discernible |
| 1 | Adoption of a position/assertion |
| 2 | Assertion + 1 fact (supporting the assertion) |
| 3 | Assertion + 2 supporting facts + 1 contrary fact |
| 4 | Assertion + 2-3 supporting facts + 2 contrary facts |
| 5 | Assertion + at least 3 supporting facts + at least 3 contrary facts; potentially further elements of argument such as explanation, objection, support (one or more items mentioned) |

Table 1: Levels of political judgements.

judgement always also contains elements of subject knowledge. Judgement competency based on subject knowledge always relates to the completion of classroom tasks with the help of the terminology or domain-specific concepts mentioned in the model. This contextualization, like valuation in the context of political issues, places particular demands on pupils (Manzel and Weisseno 2017: 66).

The model of political knowledge makes it possible to give operationalizable information about the descriptive correctness of the factual statements and normative political patterns of explanation. It contains the political values which are regarded as cognitively available subject contents, and the conceptual subject knowledge that needs to be acquired. The normative criteria are described in the domain-specific concepts relating to the basic values of the German constitution, such as peace, freedom and sustainability. This also applies to the scope of the subject knowledge which is required and must be learnt, and which is described with the concepts.

This will be described in concrete terms using the example of the construction of a test item (testlet). In the judgement test, terms related to the model of political knowledge (underlined in the text) are used to describe the initial situation (Figure 2). In this way, school-based political knowledge (according to the standards of the model) is applied in the task specification for an essay to be written by pupils.

SAMPLE TASK: DOES A PARTY BAN ON THE NPD CONTRAVENE THE DEMOCRACY PRINCIPLE?

The National Democratic Party of Germany (NPD) is hostile to foreigners and to democracy. The governments of the German Länder want to reduce the influence of the NPD. The NPD also damages Germany's reputation abroad. The interior ministers of the Länder are in favour of a ban on the NPD by the Federal Constitutional Court. After a ban, the party will no longer receive any state funding. The NPD will no longer be able to disseminate its ideas so easily. Furthermore, the terrorist NSU (National Socialist Underground) was closely allied to the NPD.

The Federal Minister of Justice (FDP) is opposed to initiating proceedings against the NPD at the Federal Constitutional Court. The members of the NPD will be able to found a replacement party after a ban. The minister does not wish to draw attention to the NPD with court proceedings. Germany is an established democracy and must deal with the NPD politically.

TASK

What should happen? Weigh up the arguments mentioned in the text for and against proceedings to ban the NPD in the Federal Constitutional Court. At the end, say which viewpoint you support and justify your decision. Use the information given in the text, not your own.

Figure 2: Sample item from the judgement test.

The task uses several of the political terms mentioned above (underlined in the text). This could give rise to certain political ideas in the minds of the students. They can relate their conceptual knowledge to the situation described in the text, that is, an intended party ban of the right-wing populist ‘National Democratic Party of Germany’ (NPD). The headline’s reference to the principle of democracy gives them orientation when looking for political arguments. There are three pro arguments and three con arguments in the text. The ‘xenophobic and anti-democratic’ party, ‘the German reputation abroad’, and the proximity to a terrorist organization (‘National Socialist Underground’) speak in favour of the ban. The cons include the foundation of a ‘substitute party’, the ‘attention through a court case’ and the necessity to ‘deal with the NPD politically’.

The following student solution is intended to illustrate the coding process:

Of course the NPD is a right-wing party and does not cast a good light on Germany, especially because of the Hitler era. But as mentioned in the text above, the members of the party could simply form a substitute party, and the whole thing starts all over again. Despite the fact that the NPD is hostile to foreigners and democracy, I think that Germany has learned from the last century and that we should deal with the NPD politically.

The student argues, weighing up the pros and cons from the text. He/She clearly takes a position (‘I think’). Two arguments in favour of a ban procedure (‘does not cast a good light’; ‘hostile to foreigners and democracy’) and two arguments against it (‘substitute party’; ‘political debate’) are presented in the correct context. The pro and contra arguments alternate and are marked linguistically with an objection (‘but’) and support (‘because of’; ‘that’; ‘of course’). Further everyday and historical explanations (‘Hitler era’; ‘past century’; ‘the whole thing starts all over again’) are inserted in the text to supplement the statements. They are not included in the evaluation because they do not relate to politics. They are not domain-specific. This response is to be classified as level 4.

MEASUREMENTS ON POLITICAL JUDGEMENT

Robust empirical findings on the capacity for political judgement are not especially numerous. Thanks to theory-based development and testing of the items, a number of studies within political didactics on subject knowledge have been able to achieve a considerable breadth of content in terms of the items included (covering the political knowledge set out in the model), with good to acceptable measurement properties. There are, however, very few robust studies on political judgement within political didactics. A preliminary exploratory study by Manzel (2007: 242 ff.) points to problems with judgement and argument. It has not been possible to sufficiently foster the skills of interpretation, deliberation and judgement in everyday teaching. The intervention also showed, however, that political judgement can be facilitated by the problem-solving strategy of identifying the pros and cons of a political problem, followed by valuation, taking into consideration individual and collective interests and values. It seems that political judgements can be stimulated by tasks that reflect controversies.

Like Osborne et al. (2004) Gronostay focuses on oral argument in the classroom in her studies. In three substudies, she has been able to show that in classes without any training in argument, the main strategy that can be observed is counterargument (indirect criticism), while objections (direct criticism) or contradictions (unsubstantiated rejection) are less common (2019a). In her studies, around a quarter of the spoken contributions were externalized arguments for and against. Only around one in ten speech acts involved the integration of criticism. Even more rarely did pupils arguing along the same lines co-construct an argument. In the experimental classes with argument training, the three high-quality speech acts, elaboration, objection and integration, were used comparatively more often. Nonetheless, the argumentative discourse of these groups was also mainly characterized by speech acts of opposition and the externalization of arguments. Here too, pupils made no concessions and expressed no explicit agreement with contributions from the opposition. She sees this as confirming the influence of the moderately competitive target structure on the nature of the speech acts used (Gronostay 2019b). She does not, however, investigate any individual learning outcomes as a result of the class discussions. Data collected by video-recording the lessons do not describe individual learning outcomes.

QUESTIONS

This article proposes an operationalizable model of judgement, which bases its graduations on a difficulty-determining quality – complexity. The judgements in the pupils' answers can be considered to be dependent on one another in their levels. This should make it possible to diagnose an increasing complexity of judgement.

The aim of this article is to show whether the individual judgements can be represented with the judgement test developed on the basis of the judgement competency model. If the model's validity is successfully proven, it can be shown that a valid test is present. The measurement model will then provide diagnostic information about political judgements. Starting from this information, interpretations and valuations can be carried out.

Against this background, this article investigates questions about the quality of the judgement test. The first question to be examined is whether the items of the test homogeneously address a judgement scale, and whether the test captures stable and widely spread differences in judgement among pupils at secondary level. The second question explores the question of validity in greater depth. The aim is to verify whether the test can only be solved with the construct of political judgement competency.

METHODS

Study design and sample

The respondents were 402 pupils in years 11–12 at a *Gymnasium* (academically oriented secondary school). Nearly equal numbers of male and female pupils took part (49.9%: 50.1%). They were presented with a 45-minute test of judgement, consisting of six text-based tasks describing open-ended political events/issues. The pupils were asked to adopt a personal position with regard to the issue, weighing up the arguments in an essay. The responses were coded in multiple stages. Participation in the study was voluntary for

the pupils. The total duration of all the tests was 90 minutes. The test was conducted by trained test administrators in 27 courses.

Measuring instruments for measurement invariance

To verify construct validity, measuring instruments for basic cognitive skills and a test of political subject knowledge were used.

Basic cognitive skills were surveyed in year levels 11 and 12 with a test from the Institute for Quality Development in Education (IQB), the Berlin test for assessing fluid and crystallized intelligence (BEFKI) (Wilhelm and Schipolowski 2012). In this study, only fluid intelligence was measured, using a subscale for deductive thinking related to figural content (figural part B, consisting of sixteen tasks). In contrast to competence, fluid intelligence is suitable because it represents a performance disposition which is largely independent of prior knowledge. According to Wilhelm and Schipolowski (2012: 17; Schroeders et al. 2015) figural-spatial tests have proven particularly valuable for the assessment of fluid intelligence, because the use of abstract drawings means that they are only slightly influenced by prior knowledge and linguistic ability.

A political knowledge test with forced-choice items was also used as a testing tool. The pupils in the university-track secondary-school level received multiple-choice tasks with four possible responses, including one correct answer. The response data were coded dichotomously. The items on school pupils' political knowledge were compiled partly from the item pool of the POWIS study (Goll et al. 2010), and partly from newly constructed items. The items were piloted several times (Figure 3). The validity of the knowledge test was verified on the one hand with an expert review, and on the other hand with calculations of correlations with the above-mentioned basic cognitive abilities of the respondents.

This study chose a new approach by measuring the levels of individual political judgements together with political knowledge. To do so, it used the model of judgement competence as its theoretical basis, thereby allowing for a differentiation between subject descriptive and subject normative knowledge as well as between subject knowledge and judgements.

The German government wants to promote sustainability by introducing a standard car toll on motorways for everyone. The political intent is that...

- Citizens switch to public transport and carpooling.
- The federal and local roads have less traffic.
- Citizens book fewer flights and industrial goods are transported more often by rail.
- Consumers demand more fuel-efficient cars from the industry and the state saves administrative costs.

Figure 3: Sample item from the knowledge test.

Scaling and model validity

The response patterns captured by a test are linked with the judgement structures formulated in the judgement model via a measurement model based on item response theory. In the judgement model, the assumption is made that more competent individuals can reach a higher level of political judgement, by explaining more complex factual contexts (complexity) in response to the situation. The mapping of a graduated response process is undertaken by means of the measuring instrument, by the specification of a stimulus (testlet) – in this case the political issue. Interdependent subtasks must be completed. In order to reach a higher level of complexity (more difficult task), all previous tasks must be completed. Because of the assumed dependencies of individual levels (tasks) in a testlet, the test is scaled with a testlet model from item response theory (cf. Wang and Wilson 2005). In a testlet model (dimension), the response – in contrast to a standard model from item response theory – is not modelled as independent; instead the response process within a testlet is modelled as dependent. For this, the testlet is understood in each case as one dimension in a within-multidimensional model. The estimation of the Rasch testlet model is carried out by fixing the covariances among the dimensions at zero, and without adjusting the difficulties on the dimensions (Brandt 2008). The identification of the testlet model is based on the mean values of the person parameters. To avoid distorting the estimation of the person parameters, the weighting of the response categories (levels of judgement) is adjusted (Haberhorn et al. 2016).

Whether the aim of the test (i.e. reliable, widely spread individual differences in judgement) is achieved is analysed by way of the variance of the person parameters. To determine reliability, the EAP/PV reliability value of the testlet scaling is considered. Before the judgement values generated by the testlet model can be considered, the validity of the testlet model must be proven. Usually the fit of the test with the measurement model is verified by way of the discriminatory power of the items, using the weighted mean square (wMNSQ) as a measure of item fit. Here it must be borne in mind that this measure of item-fit tests discriminatory power under the assumption of the stochastic independence of the response process. The significance of the item-fit findings is limited as a result. Hence the decision has been made to forego in-depth analysis of the item-fit values. The validity of the model is determined using the testlet model, by comparing the variances of the testlets with the total variance of the test. If there are higher variances on the testlets in comparison to the variance of the test as a whole, then the response processes within a testlet are dependent on one another. Two trained raters allocated the responses to the argumentation level achieved (interrater reliability $r_{\tau} = 0.89$).

Whether the homogeneous measurement values are also independent of measuring instruments is tested within the framework of construct validity. The homogeneity of the items to measure one and the same construct is supported by using latent correlations in a between-multidimensional model (Adams et al. 1997) to test the operationalization of a latent construct via the connection to other constructs. In construct validation, the construct to be measured – political judgement competency – is correlated with related or discriminating constructs. The correlations can be used to test the assumed dependencies and independencies among the constructs. This is done by means of latent correlations in a between-multidimensional model. For the test of homogeneity, the connection of judgement competency to political knowledge and fluid

intelligence – the capacity for deductive inference – is determined. Political knowledge is expected to support the connection between judgement and subject knowledge, while the capacity for deductive inference is understood as a discriminating construct. Political judgements cannot be made by means of logical inferences.

The between-multidimensional model is set up in each dimension with twenty items from the figural test, knowledge test and judgement test. The three dimensions were identified by zeroing the mean of the item difficulty in each dimension. The between multi-dimensional model provides latent correlation of the dimensions with each other, indicating whether the tests (dimensions) represent a comparable or separate construct. A high correlation means the dimensions are closely connected, and a low correlation means independence of the dimensions.

The latent correlations in the dimensions are reported. Based on the variances in the three dimensions measured by the between multidimensional model, the significance of the correlation differences can also be reported using the measure Cohen's *d*.

QUALITY OF THE TEST

For the validity of the testlet model, the variances of the testlets in comparison to the variance of the test as a whole can be considered. The first thing that becomes apparent (Table 2) is that all testlets show significantly higher variances.

One testlet shows particularly strong dependent multilevel processes. This provides a preliminary indication that the assumed dependent multilevel process is valid. In the first and last testlet, there is no empirical evidence of the assumed multilevel process. In total, only three empirically separable levels (0–2) could be formed. Other studies start with the level 1. The wMNSQ values are unremarkable and lie in an interval from 0.95 and 1.06. The EAP/PV reliability of 0.61 is low, due to the small number of testlets, the difficulty of the test, the size of the sample and the testlet model. However, the test was able to achieve a total variance of 0.78, which can be classed as good. The

| Designations | Parameters |
|---|-------------------|
| Sample size in N | 401 |
| Reliability (EAP/PV) | 0.61 |
| Total variance of the person parameters (in logits) | 0.78 |
| Discriminatory power (wMNSQ) | from 0.95 to 1.06 |
| Testlet variance (in logits) 1 | 0.79 |
| Testlet variance (in logits) 2 | 0.95 |
| Testlet variance (in logits) 3 | 0.97 |
| Testlet variance (in logits) 4 | 0.98 |
| Testlet variance (in logits) 5 | 2.44 |
| Testlet variance (in logits) 6 | 0.81 |

Table 2: Psychometric parameters of the political judgement test and the testlet model.

results of the tests of model validity and measurement invariance indicate that the testlet model can be assumed to have an adequate fit.

The analysis of construct validity shows a latent correlation of the fluid intelligence of the respondents with their political knowledge ($r = 0.28$) and judgement competency ($r = 0.23$). The difference in the correlation – measured with Cohen's d – is not significant, as $d = 0.10$. However, both correlations are so low that the tests can be assumed to be independent. The reliabilities in the three dimensions of the between-multidimensional model are good, with an average reliability of 0.68.

The low correlations with fluid intelligence (BEFKI) suggest that political knowledge and the capacity for political judgement – in conformity with expectations – can be distinguished from fluid intelligence (capacity for deductive inference). Political judgement cannot therefore be explained by drawing logical consequences. Political judgement is not based on logical conclusions, as is assumed in normative political didactics. As the testlet model shows, political judgement includes processing facts that have been sorted according to content criteria and not logical criteria. The low correlations with fluid intelligence indicate that the tests constitute a valid assessment of political knowledge and political judgement. In the field of political education didactics, Grobshäuser and Weisseno (2020) reached similar results with year 9 pupils. The study observed correlation of $r = 0.342^{***}$ between basic cognitive abilities and political knowledge. In the field of chemistry didactics, a study by Klos et al. (2008: 312ff.) likewise observed only a moderate correlation of $r = 0.263^{***}$ between basic cognitive abilities and natural-scientific working methods, thus indicating that these constructs are empirically separable. Hier ist anzumerken, dass die Korrelationen nicht latent, sondern manifest gemessen wurden.

There is a moderate latent correlation between political knowledge and judgement competency ($r = 0.52$). This proves that elements of subject knowledge play a part in judgement. The correlation makes it justifiable to speak of an independent competence dimension of judgement competency. Political judgement is connected to acquired stores of knowledge, but is not dependent on the capacity for deductive inference.

DISCUSSION

This article has investigated the question of whether the judgement test is valid. A test was used at upper secondary level, based on the model of judgement competency. The scaling of the test and verification of model validity were carried out with a testlet measurement model from item response theory. The test of homogeneity to validate the construct was carried out by comparing the variances of the testlets with the total variance of the test, and the test of measurement invariance was based on latent correlation within the framework of construct validity.

The latent correlations show, in conformity with expectations, that political knowledge is required for political judgement, but that judgement can be separated from knowledge. The test analysed with the testlet model shows that the available results tend to confirm the hypothesis of a multilevel process. The variances of the testlets are mostly greater than the satisfactory variance of the test as a whole. The differences in variance, however, are not too great, except in the case of two testlets. The results show that the multilevel process is present for most testlets. Furthermore, a look at the completion rates shows

that a lot of pupils do not even achieve the judgement on the first level. The test provides clues to this, as it is very simple in conception. However, many pupils had difficulties producing subject-appropriate arguments.

This raises the question of how, in general, a greater dependence of the levels can be empirically shown. One possibility would be to reduce the number of levels in the measuring instrument. Another possibility would be to implement the theoretical requirement with targeted practising of political judgements in essays, or with other suitable methods of cognitive activation. The prerequisite, however, is that the pupils and teachers are systematically familiarized with the criterial requirements in judgement processes – although this does not seem to be the case so far. The criterial approaches proposed in the model of capacity for political judgement should be included in teacher training. Osborne et al. (2004: 1015) show that improvement at argumentation is possible if it is explicitly addressed and taught. More work should be done towards this in institutions for practical teacher training. This would allow better assessment of judgement processes, and a clearer demonstration of the dependent levels.

The study provides indications that the pupils are able to incorporate fact-based valuations in their political judgements. The theory-based study confirms the idea behind the above-mentioned heuristic judgement grid, but takes a simpler criterial path. With the theory of political judgement that has been expounded, the desired requirements in the above-mentioned standardized examination requirements (*Einheitliche Prüfungsanforderungen*) can now be fulfilled. While the 'Beutelsbach Consensus' stipulates that pupils' subjective political opinions must not be graded at school (Weissenro 2017b), there is now a theoretical justification for grading pupils' judgements. Theory and curricular requirements can be merged.

There is a limitation to be noted, however: on the basis of the small sample and the resulting somewhat low reliability – though with a total variance of 0.78 – the test does not allow any very precise diagnostic statements. Furthermore, this study only presents the modelling of one dimension of difficulty, complexity.

To study the capacity for political judgement, then, further questions need to be answered. Conceivable dimensions could differ in the degree of independence of a judgement. This study was oriented towards political lessons, and so did not focus on judgements made on the basis of existing subjective knowledge and subjective beliefs. The latter reflect knowledge from the media and from home rather than school. The perspective-taking required in the test served the purpose of rational deliberation. This is something politicians are routinely required to do when assessing situations, and it familiarizes pupils with a multiperspectival view of the political actors and actions. As the test tasks were context-dependent, the ability to construct mental models was required. Empirical studies on different difficulty-determining features of judgement tasks have yet to be carried out.

The study has shown that these results provide stimuli for politics teaching as well as for research. The early, self-evident diagnosis of Engelhardt, that 'the formation of pupils' judgements, and the discussion of their normally differing judgements has proven to be the most difficult problem for teaching' (1968: 42), still seems valid on the surface. The quality problem of politics teaching can now be identified more precisely, as the descriptive results of this study help to show (overview of further studies in Weissenro 2019). In practice, the multitude of models has probably not been effective, as the teachers and learners have so far

lacked suitable criteria for actually setting tasks and grading them. Now that the model of political judgement competency has had its first empirical substantiation in this study, this vulnerability of political didactics and politics teaching can be closed. The quality problems associated with fostering the capacity for political judgement with conceptual subject knowledge can be resolved.

REFERENCES

- Adams, R. J., Wilson, M. and Wang, W. (1997), 'The multidimensional random coefficients multinomial logit model', *Applied Psychological Measurement*, 21:1, pp. 1–23.
- Ajzen, I. and Fishbein, M. (1969), 'The prediction of behavioral intentions in a choice situation', *Journal of Experimental Social Psychology*, 5:4, pp. 400–16.
- De Ayala, R. J. (2009), *The Theory and Practice of Item Response Theory*, New York: The Guilford Press.
- Betsch, T. and Haberstroh, S. (2005), *Current Research on Routine Decision Making: Advances and Prospects, The Routines of Decision Making*, Mahwah, NJ: Erlbaum Associates.
- Bögeholz, S., Eggert, S., Ziese, C. and Hasselhorn, M. (2017), 'Modeling and fostering decision-making competence regarding challenging issues of sustainable development', in D. Leutner, J. Fleischer, J. Grünkorn and E. Klieme (eds), *Competence Assessment in Education: Research, Models and Instruments*, Berlin and Heidelberg: Springer, pp. 263–84.
- Bögeholz, S., Hößle, C., Langlet, J., Sander, E. and Schlüter, K. (2004), 'Bewerten – Urteilen – Entscheiden im biologischen Kontext: Modelle in der Biologiedidaktik', *ZfDN*, 10, pp. 89–115.
- Brandt, S. (2008), 'Estimation of a Rasch model including subdimensions', in M. von Davier and D. Hastedt (eds), *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, vol. 1, Princeton, NJ: IEA-ETS Research Institute, pp. 51–70.
- Breit, G. (1997), 'Fragen zur politischen Urteilsbildung', in P. Massing and G. Weißeno (eds), *Politische Urteilsbildung*, Bonn: Bundeszentrale für politische Bildung, pp. 134–55.
- Brunold, A. (2017), 'Politische Partizipation als Element der Demokratietheorie in Politikwissenschaft und Politikdidaktik', in M. Oberle and G. Weißeno (eds), *Politikwissenschaft und Politikdidaktik*, Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 139–51.
- Brunswik, E. (1955), 'Representative design and probabilistic theory in a functional psychology', *Psychological Review*, 39, pp. 184–202.
- Detjen, J., Massing, P., Richter, D. and Weißeno, G. (2012), *Politikkompetenz – ein Modell*, Wiesbaden: Springer.
- Dick, R. D. (1991), 'An empirical taxonomy of critical thinking', *Journal of Instructional Psychology*, 18:2, pp. 79–92.
- Engelhardt, R. (1968), *Urteilsbildung im Politischen Unterricht: Einübung kontroversen Denkens als Aufgabe Politischer Bildung*, Essen: Neue Deutsche Schule.
- Frederking, V., Roick, T. and Steinhauer, L. (2011), 'Literarästhetische Urteilskompetenz. Forschungsansatz und Zwischenergebnisse', in H. Bayrhuber U. Harms, B. Muszynski, B. Ralle, R. Martin, L.-H. Schön, H. J. Vollmer and H.-G. Weigand (eds), *Empirische Fundierung in den Fachdidaktiken. Fachdidaktische Forschung*, Vol. 1, Münster: Waxmann, pp. 75–94.

- Goll, T. (2008), 'Standardisierung in den EPA „Sozialkunde/Politik“. Eine Zwischenbilanz', in G. Weißeno (ed.), *Politikkompetenz: Was Unterricht zu leisten hat*, Bonn: Bundeszentrale für politische Bildung, pp. 347–59.
- Goll, T., Richter, D. Weißeno, G. and Eck, V. (2010), 'Politisches Wissen zur Demokratie von Schüler/-innen mit und ohne Migrationshintergrund (POWIS-Studie)', in G. Weißeno (ed.), *Bürgerrolle heute: Migrationshintergrund und politisches Lernen*, Bonn: Bundeszentrale für politische Bildung, pp. 21–48.
- GPJE (Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung) (2004), *Anforderungen an nationale Bildungsstandards für den Fachunterricht in der politischen Bildung an Schulen: Ein Entwurf*, Schwalbach am Taunus: Wochenschau.
- Grobshäuser, N. and Weißeno, G. (2020), 'Does political participation in adolescence promote knowledge acquisition and active citizenship?', *Education, Citizenship and Social Justice*, 15:2, <https://doi.org/10.1177%2F1746197919900153>. Accessed 20 April 2021.
- Gronostay, D. (2014), '„Ich bin dagegen, weil...“: Argumentative Lehr-Lernprozesse im Politikunterricht', in S. Manzel (ed.), *Politisch mündig werden – Politikkompetenz in der Schule aufbauen und diagnostizieren*, Opladen: Barbara Budrich, pp. 35–47.
- Gronostay, D. (2019a), *Argumentative Lehr-Lern-Prozesse im Politikunterricht. Eine Videostudie*, Wiesbaden: Springer.
- Gronostay, D. (2019b), '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', *Unterrichtswiss*, 47, pp. 117–35.
- Grosser, D. (1987), 'Schwierigkeitsgrade beim Urteilen über Politik', in S. Schiele and H. Schneider (eds), *Konsens und Dissens in der politischen Bildung*, Schwalbach: Wochenschau, pp. 165–77.
- Haberkorn, K., Pohl, S. and Carstensen, C. (2016), 'Incorporating different response formats of competence tests in an IRT model', *Psychological Test and Assessment Modeling*, 58:2, pp. 223–52.
- Hostenbach, J., Fischer, H. E., Kauertz, A., Mayer, J., Sumfleth, E. and Walpuski, M. (2011), 'Modellierung der Bewertungskompetenz in den Naturwissenschaften zur Evaluation der Nationalen Bildungsstandards', *ZfDN*, 17, pp. 261–88.
- Juchler, I. (2005), *Demokratie und politische Urteilskraft*, Schwalbach: Wochenschau.
- Jungermann, H., Pfister, H. R. and Fischer, K. (2005), *Die Psychologie der Entscheidung*, Heidelberg and Berlin: Spektrum Akademischer Verlag.
- King, P. M. and Kitchener, K. S. (2002), 'The reflective judgment model: Twenty years of research on epistemic cognition', in B. K. Hofer and P. R. Pintrich (eds), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, Mahwah, NJ: Lawrence Erlbaum, pp. 37–61.
- Klieme, E., Hartig, J. and Rauch, D. (2008), 'The concepts of competence in educational contexts', in D. Leutner, E. Klieme and J. Hartig (eds), *Assessment of Competencies in Educational Contexts. State of the Art and Future Prospects*, Göttingen: Hogrefe Publishing, pp. 3–22.
- Klos, S., Henke, C., Kieren, C., Walpuski, M. and Sumfleth, E. (2008), 'Naturwissenschaftliches Experimentieren und chemisches Fachwissen: zwei verschiedene Kompetenzen', *Zeitschrift für Pädagogik*, 54:3, pp. 304–21.

- KMK (2005), *Einheitliche Prüfungsanforderungen in der Abiturprüfung Sozialkunde/Politik*, Berlin: KMK.
- Knittel, C. (2013), *Eine Feldstudie zur Untersuchung der Förderung von Bewertungskompetenz: am Beispiel der Photovoltaik*, Freiburg: Herder.
- Kolstø, S. D. (2008), 'Science education for democratic citizenship through the use of the history of science', *Science & Education*, 17:8&9, pp. 977–97.
- Lingelbach, K.-C. (1970), 'Zum Verhältnis der „allgemeinen“ zur „besonderen“ Didaktik. Dargestellt am Beispiel der politischen Bildung', in W. Klafki, G. M. Rückriem, W. Wolf, R. Freudenstein, H. K. Beckmann, K.-C. Lingelbach, G. Iben and J. Dietrich (eds), *Erziehungswissenschaft 2: Fischer Funk-Kolleg*, Frankfurt: Fischer, pp. 115–18.
- Manzel, S. (2007), *Kompetenzzuwachs im Politikunterricht: Ergebnisse einer Interventionsstudie zum Kernkonzept Europa*, Münster: Waxmann.
- Manzel, S. and Weißeno, G. (2017), 'Modell der politischen Urteilsfähigkeit: eine Dimension der Politikkompetenz', in M. Oberle and G. Weißeno (eds), *Politikwissenschaft und Politikdidaktik*, Wiesbaden: Springer, pp. 59–86.
- Massing, P. (1997), 'Kategorien politischen Urteilens und Wege zur politischen Urteilsbildung', in P. Massing and G. Weißeno (eds), *Politische Urteilsbildung*, Bonn: Bundeszentrale für politische Bildung, pp. 115–33.
- OECD (2014), *PISA 2012 Technical Report*, <https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf>. Accessed 24 July 2018.
- Osborne, J., Erduran, S. and Simon, S. (2004), 'Enhancing the quality of argumentation in school science', *Journal of Research in Science Teaching*, 41:10, pp. 994–1020.
- Petrik, A. (2012), "'...dass die Leute sich nicht auf die faule Haut legen": Rekonstruktion des Politisierungswegs einer Schülerin von libertär-sozialistischen zu marktliberalen Argumentationsmustern im Unterricht und im problemzentrierten Interview', *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 1:1, pp. 31–60.
- Pfister, H.-R., Jungermann, H. and Fischer, K. (2017), *Die Psychologie der Entscheidung*, Berlin: Springer.
- Richter, D. (2015), 'Politisches Lernen mit und ohne Concept Maps bei Viertklässlern: eine Interventionsstudie', in G. Weißeno and C. Schelle (eds), *Empirische Forschung in gesellschaftswissenschaftlichen Fachdidaktiken*, Wiesbaden: Springer, pp. 35–71.
- Riemeier, T., von Aufschnaiter, C., Fleischhauer, J., and Rogge, C. (2012), 'Argumentationen von Schülern prozessbasiert analysieren: Ansatz, Vorgehen, Befunde und Implikationen', *ZfDn*, 18, pp. 141–80.
- Schroeders, U., Schipolowski, S. and Wilhelm, O. (2015), 'Age-related changes in the mean and covariance structure of fluid and crystallized intelligence in childhood and adolescence', *Intelligence*, 48:0, pp. 15–29.
- Trautwein, U., Bertram, C., von Borries, B., Brauch, N., Hirsch, M., Klausmeier, K., Körber, A., Kühberger, C., Meyer-Hamme, J., Merkt, M., Neureiter, H., Schwan, S., Schreiber, W., Wagner, W., Waldis, M., Werner, M., Ziegler, B. and Zuckowski, A. (2017), *Kompetenzen historischen Denkens erfassen: Konzept, Operationalisierung und Befunde des Projekts „Historical Thinking – Competencies in History“ (HiTCH)*, Münster and New York: Waxmann.
- Waghid, Y. (2005), 'Action as an educational virtue: Toward a different understanding of democratic citizenship education', *Educational Theory*, 55:3, pp. 323–43.

- Wang, W.-C. and Wilson, M. (2005), 'The Rasch testlet model', *Applied Psychological Measurement*, 29:2, pp. 126–49.
- Weißeno, G. (2015), 'Konstruktion einer politikdidaktischen Theorie', in G. Weißeno and C. Schelle (eds), *Empirische Forschung in gesellschaftswissenschaftlichen Fachdidaktiken*, Wiesbaden: Springer, pp. 3–20.
- Weißeno, G. (2017a), 'Politikdidaktische Theoriebildung: eine wissenschaftstheoretische Orientierung', in M. Oberle und G. Weißeno (eds), *Politikwissenschaft und Politikdidaktik*, Wiesbaden: Springer, pp. 1–16.
- Weißeno, G. (2017b), 'Zur Historisierung des Beutelsbacher Konsenses', in S. Frech and D. Richter (eds), *Der Beutelsbacher Konsens. Bedeutung, Wirkung, Kontroversen*, Schwalbach: Wochenschau, pp. 35–56.
- Weißeno, G. (2019), 'Guter Politikunterricht: eine Zusammenschau empirischer Ergebnisse', in M. Lotz and K. Pohl (eds), *Gesellschaft im Wandel*, Frankfurt: Wochenschau, pp. 187–95.
- Weißeno, G. (forthcoming), 'Political judgements in civics lessons: A difficult challenge', in M. Oberle (ed.), *International Perspectives on Civic Education*, Frankfurt: Wochenschau.
- Weißeno, G., Detjen, J., Juchler, I., Massing, P. and Richter, D. (2010), *Konzepte der Politik: ein Kompetenzmodell*, Bonn: Bundeszentrale für politische Bildung.
- Weisseno, G. (2016), 'Political didactics and political education in Germany', in K. J. Kennedy and A. Brunold (eds), *Regional Contexts and Citizenship Education in Asia and Europe*, London: Routledge, pp. 53–65.
- Wilhelm, O. and Schipolowski, S. (2012), 'Intelligenzdiagnostik in der pädagogischen Psychologie', in G. L. Huber (ed.), *Enzyklopädie Erziehungswissenschaft Online*, Weinheim and München: Juventa, pp. 1–41.

SUGGESTED CITATION

Weisseno, Simon and Weisseno, Georg (2021), 'Political judgement competency among upper secondary-school pupils', *Citizenship Teaching & Learning*, 16:1, pp. 7–27, doi: https://doi.org/10.1386/ctl_00035_1

CONTRIBUTOR DETAILS

Simon Weisseno has worked on many research projects. He was a research assistant at Göttingen University with research interests in empirical educational research. His interests include longitudinal analyses of political knowledge, analyses of pre-vocational training programmes, motivational predispositions and cognitive activation. He is a teacher for vocational education.

Contact: Alfons-Kern-Schule, Theaterstr. 8, D 75175, Pforzheim, Germany.
E-mail: s.weisseno@googlemail.com

 <https://orcid.org/0000-0001-9013-7403>

Georg Weisseno was professor of political science and its didactics at Karlsruhe University of Education with key interests in citizenship education, education policy, evaluation methodologies and quantitative research. He has been a civics education teacher working at secondary schools. Prior to joining Karlsruhe University of Education in 1999, Georg worked from 1995 to 1999 at the University of Wuppertal. His expertise extends to programme evaluation,

research management and methodologies. His collaborative work ranges across European and national agencies, and he has led international events, publishing articles and books on citizenship education. He served from 2002 to 2006 as chair of the German Society for Civic Education Didactics and Civic Youth and Adult Education (GPJE).

Contact: Weißenburgstr. 28, D 70180, Stuttgart, Germany.

E-mail: weisseno@ph-karlsruhe.de

 <https://orcid.org/0000-0001-8349-4525>

Simon Weisseno and Georg Weisseno have asserted their right under the Copyright, Designs and Patents Act, 1988, to be identified as the authors of this work in the format that was submitted to Intellect Ltd.
