## Output-driven citizenship education

*A focus on learning outcomes to improve the quality of citizenship education*

Hoek, L.H.M.

**Publication date**
2023

**Document Version**
Final published version

# Output-driven Citizenship Education

## A Focus on Learning Outcomes to Improve the Quality of Citizenship Education

Lianne Hoek

**Output-driven Citizenship Education**

A Focus on Learning Outcomes to Improve the Quality of Citizenship Education

Lianne Hoek


Funding:
This research project was funded by Netherlands Organisation for Scientific Research (NWO), Initiative for Education Research (NRO) and the Dutch Inspectorate of Education (projects 40.5.18540.099 and 40.5.19920.001).

Cover design:
Marlinde Hoek

Output-driven citizenship education

A focus on learning outcomes to improve the quality of citizenship education

# ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 23 november 2023, te 10.00 uur

door Lianne Hendrika Maria Hoek
geboren te Almelo

***Promotiecommissie***

| | | |
|---|---|---|
| *Promotor:* | prof. dr. A.B. Dijkstra | Universiteit van Amsterdam |
| *Copromotor:* | dr. A. Munniksma | Universiteit van Amsterdam |
| *Overige leden:* | prof. dr. J. Duarte | Universiteit van Amsterdam |
| | prof. dr. G.H.P. van der Helm | Universiteit van Amsterdam |
| | dr. H. van der Veen | Universiteit van Amsterdam |
| | prof. dr. F.P. Geijsel | TIAS |
| | prof. dr. J.G. Janmaat | University College London |
| | prof. dr. A.J. Visscher | Universiteit Twente |

Faculteit der Maatschappij- en Gedragswetenschappen

**Table of Contents**

# Chapter 1

## Introduction

> Metaphorically, democracy can be seen as a ship: the ship of democracy, i.e., the political and legal institutions of democracy, only sails stable and towards its goal if the crew, i.e., the citizens of democracy, are willing and able to keep course. (De Waal, 2019, p. 32)

Democracy is a form of government essential in societies characterised by diversity in terms of culture, political ideas, religion, and more. However, a resilient, humane and healthy democracy cannot be taken for granted (De Winter, 2004). It needs to be actively maintained by citizens who underscore democratic processes (e.g., respecting the rule of law) (Council of Europe, 2018) and fundamental democratic values (e.g., equality and non-discrimination) (Levinson, 1999). For example, even if discrimination is formally and constitutionally prohibited, for discrimination to no longer occur in society, citizens' involvement and commitment is necessary (cf. De Waal, 2019) to not only be aware of one's own actions but also to have an eye for structural roots of discrimination in the distribution of opportunities, resources and power. In that sense, it is important that citizens in a democratic society are able and enabled to critically explore situations of (in)equality and (in)justice (Wardekker, 2001; Westheimer, 2008). Competence in such civic and democratic principles helps citizens navigate in a society that can be characterised as increasingly complex, individualistic, polarised and 'super diverse' (Jennissen et al., 2018; Mattei & Broeks, 2018; Organisation for Economic Co-operation and Development, 2010; Vertovec, 2007).

Many European countries have responded to the increased attention to and importance of civic and democratic competence of citizens by setting up national legislation that positions this task with schools via citizenship education (Eurydice, 2017). Citizenship education is part of the school's socialisation function (Dijkstra et al., 2018) and involves teaching young citizens the competences needed to participate in a democratic, pluriform society (Schulz et al., 2018). Next to qualification, which focuses on preparing students for further education or participation in the economy and labour market, socialisation is considered an important function of education (Council of Europe, 2018; Dijkstra et al., 2018). As education is the one thing that all young citizens have in common, schools have the potential to provide all students, regardless of their background characteristics, the opportunity to acquire the competences (i.e., the knowledge, skills, and attitude) needed to participate in a pluriform democracy (Beane, 2013).

**Towards Output-driven Citizenship Education**

In line with the increasingly acknowledged importance of citizenship education, a growing number of studies have paid attention to the characteristics of effective citizenship education (Coopmans et al., 2020; Geboers et al., 2013; Isac et al., 2014; Wanders, Van der Veen, et al., 2020; Wanders, Dijkstra, et al., 2020). Among the key research findings are studies that point to the importance of providing students with an open classroom climate in which they can critically and respectfully debate and exchange opinions (De Schaepmeester et al., 2022a; Geboers et al., 2013; Knowles et al., 2018). Moreover, it is advocated to encounter the school as a so-called 'mini-society' (cf. Dewey, 1923) and create opportunities for students to learn and practice democracy (Isac et al., 2014; Rinnooy Kan et al., 2021) and enable students' engagement in political activities in school (Hoskins et al., 2017). Despite the increasing scientific knowledge in this field, research on how schools can effectively foster citizenship competences in students remains scarce (Coopmans et al., 2020).

In contrast, school effectiveness and improvement has been more extensively studied in the educational domain that focuses on the qualification of students (e.g., via reading and mathematics) (Creemers et al., 2022; Sammons et al., 1995). In that respect, various studies have demonstrated the positive effects of an output-driven approach on students' cognitive learning outcomes (cf. Van Geel et al., 2016; Van Kuijk et al., 2016). In an output-driven approach, sometimes also referred to as a results-oriented approach (cf. Ledoux et al., 2009), the idea is to systematically use data on students' learning outcomes to ground educational decisions and, resultingly, maximise student learning (Kippers et al., 2018; Staman et al., 2017; Visscher & Ehren, 2011). The underlying building blocks of an output-driven approach differ somewhat in order and composition but largely comply with a cyclic sequence (Kippers et al., 2018; Ledoux et al., 2009) that starts with the formulation of a vision (Hilbers et al., 2010), which is translated into concrete learning objectives (Van der Kleij et al., 2015), according to which the curriculum is designed (Leeman et al., 2020), of which the learning outcomes can be measured and evaluated to see whether the learning objectives have sufficiently been met (Kippers et al., 2018; Van der Kleij et al., 2015).

Whereas an output-driven approach received considerable attention in studies conducted in the qualification domain of education (e.g., to enhance students' reading or mathematics ability), it has not been studied in the educational domain that focuses on the socialisation of students, among which is citizenship education. The substantive empirical evidence pointing to the positive effects of an output-driven approach in the qualification domain (cf. Van Geel et al., 2016; Van Kuijk et al., 2016) makes it worthwhile to examine

whether such an approach would also be applicable in citizenship education. However, several characteristics of citizenship education make the applicability less self-evident than in subjects like reading and mathematics. These characteristics include foremost the normative notion of citizenship education, meaning that what is considered ('sufficient') competence in citizenship is normative (Eidhof et al., 2016). Whereas for reading and mathematics, there are often relatively objective and detailed guidelines based on formal regulations and academic considerations available that help schools to determine the content and outcomes to be achieved, this is less common in citizenship education. For example, in the Netherlands, there is legislation on citizenship education that provides schools with relatively broad guidelines (i.e., schools need to ensure that their citizenship education does not conflict with fundamental values of democracy), but it is up to schools to decide the exact learning objectives and curriculum in citizenship education. Another characteristic of citizenship education that makes applying an output-driven approach less self-evident is the limited availability of reliable and valid measurement instruments that schools can use to gather information on student outcomes in citizenship education (Daas et al., 2016). When examining whether an output-driven approach would be possible for citizenship education, the differences between using data to enhance students' reading and mathematics ability versus their citizenship competences must not be ignored.

**This Dissertation**

The previous section explained that democracy needs to be actively maintained by citizens who underscore and have knowledge of democratic principles and fundamental democratic values, that an appeal is made to schools to foster these competences in students, and that research on how schools can effectively do this is still scarce. The section also explained that an output-driven approach is a cyclic process in which data on students' learning outcomes are used to measure whether learning objectives have been met. An output-driven approach is well-studied in the qualification domain of education and well-known for its positive effects on learning outcomes. However, in the context of citizenship education, it has not been studied. While not ignoring the substantial differences between education regarding the qualification and socialisation of students, it is relevant to examine the role of learning outcomes in improving the quality of citizenship education. Hence, the overarching research question was: 'To what extent are standardised measurement instruments able to provide for the psychometric quality and data allowing for insight into civic learning outcomes, and how can insight into civic learning outcomes be used to improve the quality of citizenship education in primary education?' In this respect, civic learning outcomes refer to the competences (e.g.,

the knowledge, attitude, and skills) that students have at a certain point of measurement without presuming that these competences are a direct result of the school (i.e., family, peers, et cetera also play a role). The remainder of this chapter elaborates on the central constructs and context of this dissertation and ends by providing an outline of the remaining chapters.

**Citizenship Competences**

Citizenship competences are generally understood as the knowledge, attitude, and skills (young) citizens need to participate in a pluriform and democratic society (Schulz et al., 2018; Ten Dam et al., 2011). In scientific research, citizenship competences have been further conceptualised in (related but) different ways. In the International Civic and Citizenship Education Study, for example, citizenship competences are conceptualised as the knowledge, attitude, and skills in four content domains: (1) civic society and systems (i.e., citizens, state institutions, and civil institutions); (2) civic principles (i.e., equity, freedom, sense of community, and the rule of law); (3) civic participation (i.e., decision-making, influencing, and community participation); and (4) civic identities (i.e., civic self-image, and civic connectedness) (Schulz et al., 2018). In the UK, the Citizenship Education Longitudinal Study has looked at, among others, citizenship interest, interest in politics, participation, mobilisation, trust, group membership, and perceived costs and benefits of participation (Keating et al., 2010). In addition, in the Citizenship Competences Questionnaire, citizenship is conceptualised as the knowledge, attitude, and skills in four so-called social tasks: acting democratically (i.e., acceptance of and contribution to a democratic society), acting in a socially responsible manner (i.e., taking shared responsibility for the communities to which one belongs); dealing with conflicts (i.e., handling of minor situations of conflict or conflicts of interest to which the student is a party); and dealing with differences (i.e., handling of social, cultural, religious, and outward differences) (Ten Dam et al., 2011).

Within citizenship competences, most studies have distinguished knowledge, attitude, and skills. Citizenship knowledge is about having (or obtaining) knowledge, insight and understanding of democratic principles and the peaceful and non-violent co-living together with citizens who hold different beliefs, religions, cultures, appearances, sexual orientations, or otherwise (Ten Dam et al., 2011). In addition, citizenship knowledge entails having knowledge about the democratic society (such as the function of laws), the underlying democratic principles (such as that all citizens are of equal value), and the application of democratic principles (such as elections) (Munniksma et al., 2017; Schulz et al., 2018).

Citizenship attitude concerns that students are (or become) willing to live up to democratic decision-making (e.g., willing to hear everyone's opinion), have the desire to

uphold social justice and do not want to harm others or the environment, are willing to search for solutions in conflict situations, and have a positive attitude towards differences (Ten Dam et al., 2011). In addition, citizenship attitude is about what young citizens think of democratic principles (like the rule of law), their perceptions of good citizenship (such as respecting individuals with different beliefs), about underscoring that all citizens have equal rights and are of equal value, and about trust in societal institutions such as the government and the media (Munniksma et al., 2017; Schulz et al., 2018).

Citizenship skills involve that students become able to pose their own opinion and listen to those of others, take a socially just position, switch perspectives and listen to others in conflict situations, can function in unfamiliar situations and adjust to the habits of others (Ten Dam et al., 2011). In addition, citizenship skills are about being able to follow a debate on television, consume news about societal or political topics, talk about societal or political topics, or participate in civic activities such as volunteer work or demonstrations (Munniksma et al., 2017; Schulz et al., 2018). In most standardised measurement instruments, citizenship skills are measured by asking students to indicate their abilities and, as such, provide information on students' self-efficacy (Keating et al., 2010; Schulz et al., 2018; Ten Dam et al., 2011).

**Measurement Instruments**

Scholars and schools can use both standardised and unstandardised measurement instruments to capture students' citizenship competences. In a standardised measurement instrument, the same questions in the same format are generally administered among all participants. An example of a standardised measurement instrument is a test or a questionnaire consisting of the same questions and answering options for all students, designed for general use and not linked to specific circumstances. Examples of standardised questionnaires in citizenship education include the instruments used in the International Civic and Citizenship Education Study (Schulz et al., 2018) and the Citizenship Education Longitudinal Study (Keating et al., 2010), as well as the Citizenship Competences Questionnaire (Ten Dam et al., 2011). Standardised measurement instruments psychometrically ensure that the same underlying constructs are measured and that the results of all participants are comparable, which is an important advantage of such instruments. In unstandardised measurement instruments, the means of measurement is not necessarily similar for all participating students. Consequently, the results of unstandardised measurement instruments are not always comparable across participants, which marks a relevant limitation. An example of an unstandardised measurement instrument is when students are asked to construct a portfolio in

which they demonstrate their competence development in a self-chosen learning goal by providing evidence (Daas et al., 2016).

Both the questionnaire and portfolio provide insight into students' citizenship competences, and schools can potentially use both types of data to improve the quality of citizenship education. However, the studies in this dissertation predominantly shed light on the psychometrics and data of standardised measurement instruments and how they can be used to improve the quality of citizenship education. The reason for this focus is that standardised measurement instruments offer some advantages over unstandardised measurement instruments, such as that the learning outcomes from all students in a group, grade or school can be compared to those of other participating schools (or even a selection of schools which is considered comparable, for example, all schools within one school board or in one region). In addition, the use of standardised measurement instruments to gain insight into students' civic learning outcomes has proven itself in scientific research (cf. Keating et al., 2010; Schulz et al., 2018; Ten Dam et al., 2011), whereas unstandardised measurement instruments generally require more development (cf. Daas, 2019). As this dissertation is a first exploration of the role of students' civic learning outcomes in improving the quality of citizenship education, the studies focused on standardised measurement instruments as they have been more thoroughly developed and studied.

**Students in Regular and Special Needs Primary Education**

The scope of this dissertation is delineated to students in primary education who are approximately ten to twelve years old. Previous studies demonstrated that students from this age group are well capable of acquiring citizenship competences (Barrett et al., 2021; De Winter, 2004; Drisko, 1993). More specifically, Drisko (1993), for example, argued that, unlike popular thought, primary school students are well capable of discussing democracy and fundamental democratic values. To illustrate this point, he shares a classroom example in which shared responsibility was fostered in students aged approximately six years old when they discussed rules and behaviour in the schoolyard, and another example in which the meaning of Jefferson's 'all men are created equal' was well discussed and interpreted with students aged approximately ten years old.

Within primary education, the studies in this dissertation were focused on students attending grades 5 and 6 of regular primary education and students attending the last ('leaving')

grade of special needs primary education[1]. Whereas these two groups are part of the primary education system in the Netherlands, they differ in student population. In this respect, students in special needs primary education often experience learning and behavioural difficulties such as externalising problem behaviour; internalising problem behaviour; a lack of motivation, effort and concentration; a slow pace of work; having a speech, language or numeracy disorder; being gifted; having an autism spectrum disorder; or being behind in literacy and or numeracy (Van der Veen et al., 2010). Recently, two studies have been conducted in the Netherlands to describe the status quo of citizenship competences of these students. In regular primary education, this measurement was carried out in 2019-2020 among students in grades 5 and 6 (approximately ten to twelve years old) (Inspectorate of Education, 2022b; Slijkhuis et al., 2021). The main findings include, among others, large differences between students based on their estimated educational attainment in citizenship knowledge and self-estimated skills, a decline in citizenship knowledge as opposed to a comparable measurement in 2009 among students in grade 6, and citizenship attitudes that are best described as 'moderately positive', for example towards cooperating in a group, dealing with differences and trust in institutions. Also, this study found that even though most schools underscore the importance of citizenship education, the formulation of a vision, learning objectives and curriculum plan were largely missing (Inspectorate of Education, 2022b; Slijkhuis et al., 2021). Even though these are valuable research findings, research on how schools for regular primary education can effectively foster citizenship competences among students remains scarce. In addition, the fact that a vision, learning objectives and a curriculum plan, as building blocks of an output-driven approach, were largely missing in schools for regular primary education illustrates an opening for improvement.

The same study design was carried out among students with learning and behavioural difficulties who attended the last ('leaving') grade of special needs primary education during the school year of 2020-2021 (aged approximately ten to twelve) (Inspectorate of Education, 2023; Slijkhuis et al., 2023). The main findings include that students in special needs primary education have less citizenship knowledge than students in regular primary education, that differences in citizenship knowledge and applied skills between students can partly be

[1] In the Netherlands, special education is divided in four domains: (1) for blind and visually impaired students; (2) for deaf and hearing impaired students and students with a language development disorder; (3) for students who are physically and mentally impaired and students who are long-term sick; and (4) for students with severe learning and behavioural difficulties (Van der Veen et al., 2010). This dissertation did not include students in special needs primary education who are blind or visually impaired, deaf or hearing impaired, physically and mentally impaired, and long-term sick students.

attributed to the classroom they are in, and that citizenship education in special needs primary education is predominantly visible in the school and classroom climate and in discussing current affairs and news. In addition, citizenship education in special needs primary education appeared to be a little less organised and concretised than regular primary education, for example, via a shared vision, concrete learning objectives and a curriculum plan (Inspectorate of Education, 2023; Slijkhuis et al., 2023). With the exception of this recent study, research into the quality, effectiveness, and learning outcomes of citizenship education in special needs primary education is largely missing. At the same time, the learning and behavioural difficulties of students in special needs primary education may pose an extra challenge in acquiring citizenship competences. This underscores the importance of research into the learning outcomes and characteristics of citizenship education in special needs primary education. In this respect, this study set out an important foundation by demonstrating that standardised measurement instruments can be used to obtain such insights.

**Context of Citizenship Education in the Netherlands**

This section provides a detailed description of citizenship education in the Netherlands as the context in which the studies in this dissertation have been conducted. Following concerns about decreased social cohesion (De Groot et al., 2022), the Dutch parliament accepted the 'Active citizenship and social integration' law in 2005. By doing so, the Netherlands was the last European country to formalise citizenship education in national legislation (Eurydice, 2005). This was no coincidence, given that the Dutch school system is characterised by high school autonomy (Dijkstra et al., 2021). As a result of this legislation, government-funded schools in the Netherlands were obligated to pay attention to the fostering of active citizenship (i.e., the willingness and ability to participate in a community and to contribute to it actively) and social integration (i.e., social participation, participation to society and its institutions, and familiarity with and engagement in expressions of Dutch culture) in the curriculum (De Groot et al., 2022; Dijkstra et al., 2021). This included regular primary education and special needs primary education (consisting of students aged approximately five to twelve), regular secondary education and special needs secondary education (consisting of students aged approximately thirteen to eighteen).

In 2009, the quality of citizenship education in regular primary education in the Netherlands was first assessed (Wagenaar et al., 2011). The findings concluded that there was a discrepancy between the desired and accomplished level of students' citizenship knowledge. In the same year, schools for regular secondary education in the Netherlands also participated in the ICCS 2009 study (Maslowski et al., 2012; Schulz et al., 2009). The research findings

pointed out that, compared to surrounding and comparable countries, Dutch students in grade 8 (approximately fourteen years old) generally had less citizenship knowledge and less interest in political and societal topics. In addition, compared to other countries, Dutch students were more reluctant towards support for equal rights for immigrants (Maslowski et al., 2012). Seven years later, the results of the Netherlands' participation in the ICCS 2016 study were presented (Munniksma et al., 2017). The main findings include that, as compared to students in comparable countries (i.e., Belgium (Flanders), Denmark, Sweden, Norway, and Finland), students in grade 8 in the Netherlands obtained lower scores on, among other things, citizenship knowledge, voting intentions, participating in peaceful demonstrations, and on trust in political parties, the parliament and other individuals in general.

In line with the findings regarding the citizenship competences of Dutch students in the previous studies, the Dutch Inspectorate of Education has repeatedly reported that the quality of citizenship education and, resultingly, citizenship competences of students in the Netherlands have ceased to develop (Inspectorate of Education, 2012, 2014, 2016, 2019, 2022a). The fact that students' citizenship competences were not at the desired level, and the assumption that a stronger focus on concrete learning objectives and learning outcomes in citizenship education would contribute to the quality of citizenship education, were among the main reasons to ground a revision of the 'Active citizenship and social integration' law from 2005. Hence, in 2021, the law 'Clarification of the citizenship education assignment to governmentally funded schools' was established (Ministry of Education, Culture and Science, 2021). The revised law prescribes that schools must visibly pay attention to a democratic and societal component of citizenship education, a social component of citizenship education, and fundamental democratic values. In addition, schools are now (i.e., since the introduction of the revised law) obligated to monitor the learning outcomes in citizenship education, including documenting to what extent they match the learning objectives.

In sum, the context in which the studies in this dissertation took place is characterised by a discrepancy between the desired and accomplished level of students' citizenship competences which seems quite robust throughout the years and manifests across different educational sectors. This illustrates the need to further improve the quality of citizenship education in the Netherlands. Moreover, the tension between a high level of school autonomy on the one hand, and the government's responsibility for the quality of (citizenship) education on the other hand, is deeply embedded in the history and culture of the Netherlands (Dijkstra et al., 2021), and manifests in legislation which is increasingly prescriptive but still formulated in rather general terms that leave much autonomy with schools. This underscores the potential

of supporting schools with empirical insights that help them to improve the quality of citizenship education, for example, by exploring the role of students' civic learning outcomes.

**Dissertation Outline**

The overarching research question of this dissertation was to examine the extent to which standardised measurement instruments are able to provide for the psychometric quality and data allowing for insight into civic learning outcomes, and how insight into civic learning outcomes can be used to improve the quality of citizenship education in primary education. To this end, four studies were conducted.

The *second chapter* focuses on measurement invariance as one of the psychometric standards of measurement instruments that received less scientific attention than other psychometric standards. An assessment of measurement invariance provides information on whether a construct is perceived and measured similarly across groups and whether the results of the measurement instrument can be used for cross-group comparisons. In this study, an assessment of measurement invariance was performed across sex, socioeconomic position, and migration background. This can be illustrated as follows: when students with a low and high socioeconomic position are asked about whether they want to contribute to a classroom conversation about the news, differences in access to news between these two groups of students (e.g., via a newspaper subscription or personal mobile phone) may cause systematic differences in how these two groups answer this question. Instead of measuring the construct of 'democratic attitude' in both groups, this question may measure 'access to news' among students with a low socioeconomic position. In addition, differences in access to news may have increased throughout the years (i.e., students with a mobile phone nowadays have access to a wealth of information whereas those without, do not), causing systematic differences in how groups respond to this question that were not present at an earlier point in time. This illustrates that it is important to periodically assess measurement invariance, particularly when it entails instruments that measure dynamic constructs such as citizenship education that take on meaning in society (Mattei & Broeks, 2018) and of which the meaning may change over time (Putnick & Bornstein, 2016). This study provides the findings of an assessment of measurement invariance using a standardised measurement instrument for citizenship competences that was developed over a decade ago.

The *third chapter* uses a standardised measurement instrument to describe the characteristics and citizenship competences of students in regular primary education and special needs primary education. In addition, this study examined whether the relationship between attending special needs primary education and citizenship competences changed under

the influence of various characteristics of citizenship education. Studies conducted in schools for regular primary education have, for example, demonstrated that positive teacher-student and student-student relationships promoted learning outcomes in general (Baker et al., 2008; Hamre & Pianta, 2001; Roorda et al., 2021; Wentzel, 2017) and in the context of citizenship education (Sampermans et al., 2018; Wanders, Dijkstra, et al., 2020; Wanders, Van der Veen, et al., 2020). Other studies have demonstrated that students with learning and behavioural difficulties are more at risk of experiencing negative relationships with their teachers and peers (Freire et al., 2020; Pinto et al., 2019; Zee & Roorda, 2018). As such, it may be the case that the relationship between special needs primary education and citizenship competences is more positive for students who experience more positive teacher-student and student-student relationships. This study is one of the first empirical studies to provide insight into the citizenship competences and characteristics of citizenship education among students in special needs primary education.

The *fourth chapter* described the development of the Fundamental Democratic Values Questionnaire. By doing so, it elaborated on its ability to provide for the psychometric quality and data, allowing insight into young students' attitudes towards fundamental democratic values. Underscoring fundamental democratic values such as equality, non-discrimination, and tolerance is essential for the peaceful co-living of diverse (groups of) individuals in a pluriform and democratic society (Dekker & Den Ridder, 2016; Levinson, 1999; Welzel & Dalton, 2014). In this respect, the widespread legitimation of fundamental democratic values is visible in national constitutions and international policy documents such as the Universal Declaration of Human Rights (UDHR), the European Convention on Human Rights (ECHR), the United Nations Convention on the Rights of the Child (UNCRC) and the Reference Framework of Competences for Democratic Culture (RFCDC). In addition, citizens, too, largely agree that it is important to teach fundamental democratic values in schools (Dekker & Den Ridder, 2016; Van Goethem et al., 2020). Schools are given an important role in fostering fundamental democratic values among young students via citizenship education (Beane, 2013; Council of Europe, 2018; De Winter, 2004; White, 1999). In this respect, the Fundamental Democratic Values Questionnaire not only provides valuable empirical information on young students' attitude towards values essential for the perseverance of democracy, but also provides schools with important insight into civic learning outcomes, which can be used to improve the quality of citizenship education further.

The *fifth chapter* constructs a model for output-driven citizenship education and theoretically reflects on the feasibility of implementing such an approach in the context of

citizenship education. Whereas (aspects of) an output-driven approach have been well-studied in the qualification domain of education (Van Geel et al., 2016; Van Kuijk et al., 2016), it has not been studied in the context of citizenship education (Eurydice, 2017). The model for output-driven citizenship education consists of five building blocks which have been derived from the literature focused on learning outcomes in the qualification domain of education (cf. Kippers et al., 2018; Ledoux et al., 2009; Van der Kleij et al., 2015; Van Geel et al., 2016; Van Kuijk et al., 2016). The building blocks are discussed in light of the opportunity and challenges that play a role when implementing an output-driven approach to the context of citizenship education.

At last, the *sixth chapter* summarises all findings and elaborates on the implications, limitations and directions for future research.

## Chapter 2

## Measuring Citizenship Competences: Assessment of Measurement Invariance*



**Abstract**

Standardised questionnaires are used to measure the outcomes of citizenship education. These outcomes are often compared across groups to document different outcomes, for example, between boys and girls. A prerequisite for cross-group comparisons is an assessment of measurement invariance.

This study used data from 6035 students from 87 Dutch primary schools to examine the measurement invariance of the Citizenship Competences Questionnaire (Ten Dam et al., 2011). Dutch schools use this questionnaire to gain insight into students' citizenship knowledge, attitudes, and skills. Measurement invariance was assessed across sex, socioeconomic position, and migration background.

Measurement invariance was sufficient in most cases, allowing for cross-group comparisons of associations between latent constructs and their indicators, and in some cases, for cross-group comparisons of the latent means. We conclude by emphasising that periodic assessment of measurement invariance in instruments measuring citizenship competences is important due to the dynamic nature of the construct.

*Keywords:* Citizenship competences; Measurement invariance; Measurement equivalence; Citizenship education; Social outcomes.

*This chapter is based on: Hoek, L., Zijlstra, B. J. H., Munniksma, A., & Dijkstra, A. B. (2023). Measuring citizenship competences: Assessment of measurement invariance. *Journal of Social Science Education, 22(2).* https://doi.org/10.11576/jsse-5837

## Introduction

Like other educational domains, it is essential to gain insight into what students learn in citizenship education. This insight can be used to facilitate the learning process and evaluate the contents and delivery methods of the curriculum. There are several ways to obtain insight into student outcomes, including standardised measurement instruments (Daas et al., 2016). Using standardised measurement instruments is beneficial in many ways, for example, because of their practical usefulness and the opportunities for securing the quality and validity of the measurement. This is why standardised questionnaires have been widely used to measure students' citizenship competences in terms of knowledge, attitude, and skill (Ireland et al., 2006; Schulz et al., 2018).

Multi-group comparisons based on standardised questionnaires have demonstrated that citizenship competences are related to students' background characteristics, such as sex, socioeconomic position (SEP), and migration background - both internationally (Ireland et al., 2006; Kerr et al., 2007) and in the Dutch school system in which our data were gathered (Dijkstra et al., 2015; Geijsel et al., 2012). These findings are important for educational practice because they may lead to adaptations in the contents or delivery methods of citizenship education so that the learning objectives are met, and all students benefit optimally. However, a prerequisite for meaningful comparisons across different population groups (e.g., boys and girls) is that the construct to be measured is understood similarly in each group (Isac et al., 2019; Steinmetz et al., 2009). In other words: it is essential to know whether the difference in citizenship competences between, e.g., boys and girls, is a 'true' difference or a difference caused by boys and girls systematically understanding questions in a different way. To determine whether a questionnaire measures the same across groups, the measurement invariance should be assessed (Meredith, 1993).

Assessment of measurement invariance is important for the methodological quality of a measurement instrument (Meuleman et al., 2022). However, as the interpretation of a construct can change over time (Putnick & Bornstein, 2016), the establishment of measurement invariance is not a static given. This applies in particular to citizenship education, which is a dynamic construct that, like other societal phenomena, takes on meaning in and moves along with changes in society (Mattei & Broeks, 2018). This underscores the importance of periodic assessment of measurement invariance in student questionnaires, as is the case for measuring citizenship competences. Examples of such changes, amongst others, include increased cultural diversity (Eurostat, 2021; US Census Bureau, 2020) or social inequality (Organisation for

Economic Co-operation and Development, 2017). Both changes, for instance, may have influenced the public debate and, resultingly, have led to different interpretations of constructs related to citizenship education.



For example, students with and without migration backgrounds may perceive debates about how to deal with socioeconomic or socio-cultural differences differently nowadays, in times of what is sometimes referred to as a shift from diversity to 'super-diversity' (Vertovec, 2007), as opposed to earlier in time, also influencing the response in measurement instruments regarding these topics. Hence, it is important to periodically assess the measurement invariance in instruments that measure citizenship competences, in case changes in society may have shifted the norms and values of citizens regarding value-sensitive themes (Munck et al., 2018).

Whereas both conceptualisation and operationalisation are vital steps in developing measurement instruments, the assessment of measurement invariance relates explicitly to the operationalisation of a measurement instrument. It assumes that the conceptualisation of the instrument is carefully considered and, therefore, not associated with differences in student characteristics like sex, SEP, migration background, religiosity, age and others. The assumption of such a generic conceptualisation of citizenship is the basis of large-scale standardised measurement instruments for citizenship competences. National and international examples include instruments like the International Civic and Citizenship Education Study (ICCS) for international comparisons used in periodic cycles (Schulz et al., 2018), the Citizenship Education Longitudinal Study (CELS) for longitudinal research in the UK (Cleaver et al., 2006), or the Citizenship Competences Questionnaire (CCQ) for annual measurements used by schools in The Netherlands (Ten Dam et al., 2011). In the conceptualisation of such instruments, important building blocks are core democratic values, trust, participation and involvement, and the support for institutions of a democratic society. Such instruments generally move away from conceptualisations that take a position regarding an 'ideal' balance between values like, for example, individual freedom and solidarity towards others – also known as contested or school-specific goals in citizenship education (Eidhof et al., 2016). Rather, these conceptualisations adhere to consensus goals, referring to commonly shared values of reciprocity of 'treating others as you want others to treat you' – also known as *regula aurea* or the Golden Rule (cf. Etzioni, 1996; Wattles, 1996).

Despite its importance, research on measurement invariance in instruments for citizenship competences is still scarce. It is predominantly assessed in between-country research (Byrne & Watkins, 2003), such as the ICCS. The technical report of the ICCS (Schulz et al., 2016), for example, demonstrated that most of the citizenship constructs show

measurement invariance up to a level where associations of latent constructs can be safely compared across countries, but not always to the level where comparison of means of these latent constructs is justified. In addition, Isac et al. (2019) assessed the measurement invariance of a specific part of the ICCS questionnaire that focuses on young people's tolerant attitude towards immigrants (i.e., support for equal rights). The authors found that most items were measurement invariant to the level where average scale scores can be validly compared across European countries. However, the assessment of measurement invariance is also important in within-country comparisons (Steinmetz et al., 2009; Vandenberg & Lance, 2000). In this respect, some researchers mentioned that homogeneity of the population in terms of the measures being compared across groups is often implicitly assumed, regardless of whether it was assessed (Muthén, 1989; Steinmetz et al., 2009).

This study aims to examine the measurement invariance of a Dutch questionnaire for assessing citizenship competences across groups. By doing so, we assess whether this questionnaire, which was designed over fifteen years ago, is still[2] valid across different groups in society today, as societal changes may have affected how different groups interpret the questionnaire items. In this way, we contribute to the insight into the assessment of measurement invariance in within-country comparative research on citizenship competences, as this empirical knowledge base on measurement invariance in these questionnaires is still scarce. Indirectly, this study sheds light on whether the measurement of citizenship competences is robust amid changes in society and population. For educational practice, this study serves as a validity check of a measurement instrument that schools use to improve their citizenship education.

Following previous research showing relevant and robust differences in citizenship competences based on sex, SEP, and migration background (Geijsel et al., 2012; Ireland et al., 2006; Kerr et al., 2007), we focus on the assessment of measurement invariance across these three student-background characteristics. The findings of this study may underscore the results of these previous studies, which are largely built on the same generic conceptualisation of citizenship competences, in case the measurement instrument appears largely measurement invariant or, alternately, place caution on the findings of these studies in case measurement invariance could not be established. Moreover, various scholars have advocated investigating measurement invariance over sex, SEP, and migration background (Kline, 2015; Wray-Lake

[2] Still, as was the case during the initial construction phase when the measurement invariance was tested as part of a broad set of psychometric tests, meeting all necessary requirements (Ten Dam et al., 2011; Geijsel et al., 2012; based on personal communication, because these results were not available in print).

et al., 2017) or have stressed the importance of examining similarities and differences in citizenship competences based on these groups (Cleaver et al., 2006). The instrument used for testing measurement invariance is the CCQ, a large-scale standardised measurement instrument based on a generic conceptualisation of citizenship competence. The CCQ is used primarily for annual measurements of students' citizenship competences at Dutch schools. The questionnaire has been used for over a decade – yielding a rich dataset of consecutive cohorts that is eminently suitable to assess whether contextual changes have deteriorated alignment between the instrument and its context.



## Theory

### Meaningful Comparison of Groups

In this section, building on He and Van de Vijver (2013) and Isac et al. (2019), we outline three examples of how the operationalisation of measurement does not measure the same across groups. We also provide a 'counter-example' of a difference between groups that is not the result of measurement non-invariance but an example of a 'true' difference. To illustrate our examples, we used existing items from the CCQ (measuring citizenship attitudes, skills, and knowledge of students). Students needed to indicate to what extent the items applied to them on a four-point Likert scale ranging from 'does not apply at all to me' to 'applies completely to me' for attitude items or ranging from 'not good at all' to 'very good' for skill-items. Students had to pick the best answer in the knowledge items by choosing one out of three. The use of these items at this place is merely illustrative.

First, we outline how measurement non-invariance is caused by the fact that some underlying items are not considered indicative of the construct to be measured for some group members. To illustrate this for SEP, we look at the item: 'If we talk about the news in class, I want to add something to the conversation too'. Compared to students with a high SEP, students with a low SEP may have limited access to news sources, such as a subscription to a daily newspaper or a personal mobile device to check news websites. Therefore, students with a low SEP may answer this item with 'does not apply at all to me', whereas they are willing to contribute to the conversation. Therefore, their answer indicates their accessibility to news, not their democratic attitude. This may cause students with a low SEP to be wrongly labelled as 'less capable' of the construct 'democratic attitude' because this specific indicator of the construct is less applicable to their context.

Second, we outline how measurement non-invariance is caused by some group members understanding items differently due to linguistic differences. To illustrate this for migration

background, we look at the item: 'How good are you at… holding on to your opinion, if you are really right?' Linguistic differences between students with and without a migration background may influence how students understand constructs or underlying items. Language that involves ambiguity in meaning (e.g., 'holding on to something', 'if you are really right') or metaphorical language is particularly susceptible to differences in interpretation, for example, by students who have another mother language (more likely being students with a migration background). However, vice versa, multilingual students (more likely students with a migration background) may also benefit from their knowledge and skill in language comprehension. This may be less common for monolingual students (more likely students without a migration background). Either way, such linguistic differences may distort the results and wrongly label students as more or less skilled in citizenship competences.

Likewise, group members may understand items differently due to cultural differences. To illustrate this for sex, we look at the item: 'It is normal to help in the household (for example, by preparing the dinner table, tidying up or cleaning)'. In some cultures, helping in the household is considered predominantly a task for girls and not for boys. Boys who grow up in such a culture may be more likely to answer this question with 'does not apply to me at all', whereas their attitude towards 'acting in a socially responsible manner' may, in reality, be different than this answer reveals. Or vice versa: girls who grow up in such a culture may be more likely to answer this question with 'applies completely to me', whereas their attitude towards 'acting in a socially responsible manner' may, in reality, be different than this answer reveals.

Third, we outline how measurement non-invariance is caused by some groups characterised by a specific response style. To illustrate this for sex, we look at the item: 'People who earn sufficient salary should together take care of people with less wealth'. Suppose that girls are more likely to answer items in a more socially desirable way and systematically answer this item with 'applies completely to me', whereas boys answer this item in a less socially desirable way. This may cause girls to overestimate their attitude toward the construct 'acting in a socially responsible manner'.

At last, we outline how a difference can be not the result of measurement non-invariance but a 'true' difference between groups. To illustrate this, we look at the item: 'In a sports game, the referee takes a wrong decision *against* your team. What should you do?' And the following answer options: (a) Go to the referee and debate the decision; (b) Get the coach of your sports team; (c) Keep on playing because the decision of the referee is directive during the game. The latter is appointed the preferred answer. Student background characteristics such as sex, SEP,

or migration background may influence how group members respond to this item. However, these differences are not an example of measurement non-invariance when the differences do not adhere to differences in understanding, interpretation, or applicability to the context, but rather are an example of 'true' differences in what is valued about acting in a socially responsible manner. Regardless of these 'true' differences, the conceptualisation of acting in a socially responsible manner holds that if participants of a sports game agree on rules beforehand, they conform to these rules during the sports game – even if they do not agree during the game.



**Assessment of Measurement Invariance**

Multiple-group confirmatory factor analysis (MGCFA) is the most commonly used method to assess measurement invariance (Putnick & Bornstein, 2016). In MGCFA, hierarchical, subsequent models with increasing restrictions are specified and compared. These are the configural model, the metric model, and the scalar model.

The *configural model* assesses whether the instrument measures the same latent factors across groups and whether the indicators are the same across groups (Isac et al., 2019). To test for configural invariance, models with the same pattern (i.e., the same configuration) should be specified across groups (Vandenberg & Lance, 2000), meaning an equal number of latent variables, indicators, et cetera. If the configural model yields poor model fit, it indicates that in one of the groups, a different pattern fits the data. For example, for boys, one of the questions may not be an indicator of the latent construct, whereas, for girls, it is. If the configural model fits the data well, it provides ground for testing metric invariance.

The *metric model* assesses whether the factor loadings differ across groups (Horn & Mcardle, 1992). To test for metric invariance, the factor loadings are constrained to equality across groups (Reeskens & Hooghe, 2010). Thus, for example, two equal path models with the same pattern are specified with the first factor loading in the model for boys constrained to equality to the first factor loading in the model for girls, and so on for the remaining factor loadings. If the metric model yields poor model fit, it indicates that in one of the groups, a factor loading relates differently to the latent variable as compared to the other group (e.g., for boys, the third item is strongly related to the latent variable, but for girls, this is not the case). If the metric model fits the data well, it allows for subsequent analyses of testing scalar invariance. Reaching metric invariance justifies the comparison of latent variables and their associations across groups (Isac et al., 2019). In addition, achieving (partial) metric invariance is seen as a minimal prerequisite for meaningful cross-group comparisons (Little, 2013).

The *scalar model* assesses whether the intercepts (i.e., the constant or the scalar) of the

indicators are the same across groups (Steenkamp & Baumgartner, 1998). To test for scalar invariance, the intercepts of the indicators are constrained to equality across groups. Thus, in addition to the identical pattern (configural model) and equal factor loadings (metric model) across groups, the intercepts are constrained to equality. This means that the intercept of the first item for, e.g., boys is constrained to equality to the intercept of the first item for girls, and so on for the intercepts of the remaining items. If the scalar model yields poor fit, at least one intercept differs across groups (Isac et al., 2019). If the scalar model fits the data well, it justifies comparing latent means across groups (Reeskens & Hooghe, 2010).

## Methodology

In this study, we examined the measurement invariance in the assessment of citizenship competences (i.e., the competence of students to function and participate in society) across sex, SEP, and migration background. We did so by looking at competences of students in terms of attitude (i.e., thoughts, desires, and willingness), knowledge (i.e., knowing, understanding, insight), and skill (i.e., an estimate by students of what they think they are able to).

### Data

We used data of consecutive cohorts from the CCQ (Ten Dam et al., 2011). Schools use the CCQ to measure students' citizenship knowledge, attitude and skill in terms of four so-called 'social tasks': acting democratically, acting socially responsible, dealing with conflicts, and dealing with differences (Ten Dam et al., 2011). The CCQ is suitable for grades 5 and 6 of primary education (approximate age is 10 to 12 years old) and grades 7, 8, and 9 from secondary education (approximate age is 12 to 16 years old). We retained this study's scope to data gathered in primary education for pragmatic reasons. In this sample, we merged data from grades 5 and 6 to obtain larger group sizes and more robust results.

### Sample and Procedure

The sample consists of 6035 students from 87 primary schools participating in the Dutch' Alliance Citizenship' (Table 1). The Alliance Citizenship is a partnership of schools that organises annual measurements of citizenship competences of students. This study used data from 2015, 2016, 2017, 2018, and 2019. We used a sample that was both recent and large enough to detect possible changes in society and population composition. Each year, the CCQ is online available to participating schools during spring. The sample of schools differed each year: some schools participated more than once; some only once[3].

[3] The possibility that some students were measured twice (i.e., from grade 5 in one years' measurement to grade 6 in the following years' measurement), as well as the nested structure of the data, may have caused a degree of dependency that we did not account for. In practice, it means that our results may be somewhat too conservative.

**Variables**

This section provides information on the conceptual framework underlying the questionnaire that we examined in our analyses (Table 2). The values for Cronbach's alpha are calculated with the sample used in the present study. The values were largely consistent with the original alpha's provided by Ten Dam et al. (2011), based on 16,000 students from grade 6 and grade 9 who participated in 2005, 2006, and 2007.[4] The variables Attitude – Acting democratically and Skill – Acting democratically are both presented in two interpretable factors. The variable Skill – Acting in a socially responsible manner is presented in a combined factor with the scale Skill – Dealing with conflicts.

[4] In order to only share robust results, the results of the questionnaire are reported at the overarching scale-level (e.g., knowledge, attitude, or skill), and not on subscale-level. We performed the analyses of measurement invariance on both the scale- and subscale-level to gain more information on exact items that could be hindering reaching a higher level of measurement invariance. Hence, we provided the values for Cronbach's alpha also on the subscale-level (of which some values are below the advised threshold of 0.70 (Cortina, 1993)).

**Table 1**

*Descriptive Information of the Sample*

| | 2015 | 2016 | 2017 | 2018 | 2019 | Total |
|---|---|---|---|---|---|---|
| Primary schools (N) | 26 | 18 | 16 | 11 | 19 | 87 |
| Students (N) | 1587 | 1349 | 1176 | 564 | 1359 | 6035 |
| Grade | | | | | | |
| Grade 5 | 809 | 675 | 605 | 296 | 678 | 3063 |
| Grade 6 | 778 | 674 | 571 | 268 | 681 | 2972 |
| Age | | | | | | |
| 10 years or younger | 309 | 225 | 198 | 121 | 339 | 1192 |
| 11 years | 676 | 638 | 570 | 272 | 643 | 2799 |
| 12 years | 389 | 391 | 315 | 137 | 301 | 1533 |
| 13 years | 55 | 33 | 30 | 9 | 19 | 146 |
| 14 years | 1 | 3 | 0 | 0 | 0 | 4 |
| 15 years | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 years or older | 1 | 0 | 1 | 1 | 1 | 4 |
| Sex | | | | | | |
| Boy | 734 | 642 | 555 | 254 | 628 | 2813 |
| Girl | 697 | 648 | 559 | 285 | 675 | 2864 |
| SEP | | | | | | |
| Low SEP [a] | 430 | 402 | 393 | 192 | 535 | 1952 |
| High SEP [b] | 608 | 526 | 425 | 178 | 364 | 2101 |
| Migration background | | | | | | |
| No migration background [c] | 1102 | 989 | 847 | 401 | 1086 | 4425 |
| Migration background [d] | 334 | 302 | 267 | 139 | 217 | 1259 |

*Note.* SEP = socioeconomic position. Low SEP = highest educational level of mother and father is 'no school', 'only primary education' or 'only secondary education'; high SEP = highest educational level of mother and father is 'higher education'. No migration background = both parents are born in the Netherlands; migration background = at least one parent is born outside of the Netherlands.

**Table 2**

*Conceptual Framework of Citizenship Competences*

| Components / Social tasks | **Knowledge (α = .79)** Knowing, understanding, insight<br>A young person with such knowledge… | **Attitudes (α = .89)** Thoughts, desires, willingness<br>A young person with such attitudes… | **Skills (α = .87)** Estimate of what one can do<br>A young person with such skills… |
|---|---|---|---|
| **Acting democratically**<br>Acceptance of and contribution to a democratic society | … knows what democratic principles are and what acting in accordance with them involves (8 items, α = .67) | *… wants to hear everyone's voice, enter into a dialogue* (3 items; α = .65) *and make an active, critical contribution* (3 items; α = .65) | *… is able to assert own opinions* (3 items; α = .74) *and listen to the opinions of others* (3 items; α = .68) |
| **Acting in a socially responsible manner**<br>Taking shared responsibility for the communities to which one belongs | … knows social rules (i.e. legal or unspoken rules for social interaction) (6 items, α = .54) | … wants to uphold social justice, is prepared to provide care and assistance, does not want to harm another or the environment as a result of his or her behaviour (6 items, α = .68) | … can adopt a socially just position (5 items, α = .76) |
| **Dealing with conflicts**<br>Handling of minor situations of conflict or conflicts of interest to which the child is a party | … knows methods to solve conflicts such as seeking win-win solutions, calling in help from others, admission of mistakes, prevention of escalation (7 items, α = .62) | … is willing to explore conflicts, prepared to consider the standpoint of another, jointly searches for an acceptable solution (6 items, α = .79) | … can listen to others, put oneself in someone else's place, seek win-win solutions (5 items, α = .76) |
| **Dealing with differences**<br>Handling of social, cultural, religious, and outward differences | … is familiar with cultural differences, has knowledge of rules of behaviour in different social situations, knows when one can speak of prejudice or discrimination (6 items, α = .63) | … has a desire to learn other people's opinions and lifestyles, has a positive attitude towards differences (6 items, α = .85) | … can adequately function in unfamiliar social situations, adjust to the desires or habits of others (4 items, α = .67) |

*Note.* The conceptual framework is derived from (Ten Dam et al., 2011)

For the Knowledge items, students chose the answer they thought was the best response. An example is: *All children have a right to… (a) an allowance; (b) choose who they want to live with;* or *(c) education*. The correct answer here is c. All knowledge items were dichotomised into correct (1) or incorrect (0). The phrasing of the Attitude items is: *How well does this statement apply to you?* A sample statement is: *I like knowing something about different religions*. The response options are: *(1) does not apply at all to me, (2) does not apply much to me, (3) applies a fair amount to me,* or (*4) applies completely to me*. The phrasing of the Skill items is: *How good are you at…?* A sample statement is: *… finding a solution that everyone is satisfied with for a conflict?* The response options are: *(1) not good at all; (2) not very good; (3) pretty good;* or *(4) very good.*

For sex, boys (49.55%) were appointed with value 1 and girls with value 2. For SEP, we converted maternal and paternal educational levels into a new variable, indicating whether the mother or father's highest obtained educational level is either no school, primary school, or secondary school (value 1) or higher education (value 2). In our sample, 48,16% of all students had a low SEP. For migration background, we converted maternal and paternal country of birth into a new variable, indicating whether both parents of the students were born in the Netherlands (value 1) or whether at least one parent of a student has a migration background, regardless of the country of origin (value 2). In our sample, 77,85% of all students had both parents born in the Netherlands.

**Analytic Plan**

We conducted MGCFA and compared groups based on sex (boys versus girls), SEP (low versus high), and migration background (having no migration background versus having a migration background). Analyses were conducted with R version 3.5.0 (R Core Team, 2021), using the lavaan package (Rosseel, 2012) and the semTools package (Jorgensen, 2021), with particular focus on the guidelines in the measEq.syntax. While doing so, analyses were performed according to a the following principles.

First, we treated the indicators that were directly observed (questionnaire items) as ordered categorical factors because they were measured using a four-point Likert scale (attitude and skill) or a three-option multiple-choice (knowledge). Therefore, they could not be treated as continuous indicators. Assessing measurement invariance using ordered categorical factors makes the procedure more complex because it also involves testing for threshold invariance. Testing for threshold invariance is a means to do justice to the ordered categorical nature of indicators. The threshold model assumes that a normally distributed latent item-response lies underneath each observed categorical indicator (Kite et al., 2018). The threshold can be seen

as a 'tipping point' between the different category responses (e.g., between answering 'does not apply at all to me' and 'does not apply much to me'). The threshold model is estimated before the metric model, but only if polytomous indicators are involved (thus: only for attitude and skill).



In the case of dichotomous indicators (such as the knowledge items), it is not possible to distinguish between equality constraints on the thresholds (testing threshold invariance), factor loadings (testing metric invariance), and intercepts (testing scalar invariance) (Wu & Estabrook, 2016). Instead, the equality constraints need to be added simultaneously. Therefore, we tested only for configural and scalar invariance for dichotomous indicators.

**Data Preparation**

All empty cases were dropped via list-wise deletion by removing all rows with less than nine items answered. This removed the rows in which only certain standard school characteristics (e.g., unique identifier for school, year and class) were automatically filled in (8 items) but none of the actual questionnaire items. The resulting data had 1.28% missing observations that did not show any pattern of missingness. The percentage of missing data is considered acceptable as basis for further analyses (Bennett, 2001). Second, four variables were excluded from our subsequent analyses because they consisted of three items, which was insufficient as input to our fit measures in the configural model. These excluded variables are: 'Attitude – Acting democratically 1', 'Attitude – Acting democratically 2', 'Skill – Acting democratically 1', and 'Skill – Acting democratically 2'.

**Model Selection Procedure**

To assess how well the configural, threshold, metric, and scalar models fit our data, we consulted the Chi-square test of overall model fit ($\chi^2$), and the relative difference in model fit ($\Delta \chi^2$), the comparative fit index (*CFI*), and the root mean square error of approximation (*RMSEA*) – where the CFI and RMSEA are functions from the Chi-square test statistic (Shi et al., 2019). We followed five decision rules in the model selection procedure using these model fit measures.

The first decision rule is based on the *CFI*. We consulted the *CFI* to indicate how well the model fit improved compared to the null model. The *CFI* considers the complexity of the model and ranges from 0 to 1. A value above 0.95 is preferred (Hu & Bentler, 1999).

The second decision rule is based on the *RMSEA*. The RMSEA indicates the "badness-of-fit" (Adelson, 2012) and considers the model complexity by estimating approximation error per model degree of freedom. Larger values of the *RMSEA* indicate a worse model fit. An *RMSEA* lower than 0.05 indicates 'close fit', and an *RMSEA* between 0.05 and 0.08 indicates

'satisfactory fit' (Browne & Cudeck, 1993). In this study, we perceived a value for RMSEA ≤ 0.08 as acceptable.

The third decision rule is based on the Chi-square test of overall model fit. This fit index tests the assumption of respectively configural, threshold (if applicable), metric (if applicable), and scalar invariance. A downside of the Chi-square test for overall model fit is that even minor deviations from the suggested models can reject the null hypotheses of model fit for large samples (Shi et al., 2019). Therefore, we proceeded to test additional invariance models as long as values for *CFI* and *RMSEA* of the configural model are acceptable – even if the Chi-square value is significant. At the same time, we performed permutation tests to ensure that the rejection of the overall model fit is most likely due to the same underlying reason across groups (Jorgensen, 2017). In permutation tests, all observations are randomly reassigned (a thousand times) to groups. These permuted group compositions are expected to fit the data equally poor or well as the hypothesised group composition (in which we 'manually' assigned, e.g., all boys to the one group and all girls to the other group). If this is the case, it is indicated by a non-significant p-value. However, a significant p-value suggests that there might be different underlying reasons across groups causing the significant Chi-square test of overall fit.

The fourth decision rule is based on the p-value of the Chi-square test of the difference in model fit ($\Delta \chi^2$). This fit index was applied to test the null hypothesis of equal model fit for two models. If the p-value of the Chi-square test of difference is significant, it indicates a meaningful difference between the two models in comparison, and we reject the more restricted model.

The fifth decision rule entails whether or not to test for partial invariance. A partial model can be seen as an 'in-between model' because some, but not all, of the equality constraints may need to be rereleased. To assess what constraints this involves, we inspected standardised mean residuals and modification indices. Standardised mean residuals can be consulted to identify which factor loadings, thresholds, or intercepts differ most across groups and may need to be rereleased. In addition, modification indices estimate how much the Chi-square test statistic of a model decreases if a constrained parameter is set to free again. We used a Bonferroni-adjusted alpha level to see whether modification indices were significant and could be considered released. When values for *CFI* and *RMSEA* are within the thresholds, but the p-value for the Chi-square test of difference is significant, we additionally tested for partial measurement invariance. To maintain analyses within the scope of this article, we only searched for partial invariance if it established a meaningful difference, i.e. reaching partial metric invariance (for polytomous indicators) or partial scalar invariance (for polytomous and

dichotomous indicators). We did not test for partial invariance if the meaningful minimum of partial metric invariance (Little, 2013) was still more than one step away (e.g., when the configural model fitted well, but the threshold model did not).



## Results

We specified two types of first-order one-factor models: (1) models in which social tasks such as 'acting democratically' are specified as indicators of overarching constructs such as citizenship attitude, knowledge, or skill; and, zooming in, (2) models in which items from the questionnaire are specified as indicators of the aforementioned social tasks. We summarised the findings in Tables 3, 4 and 5. In this section, we elaborate on the results by first sharing the results of the models with social tasks as indicators of citizenship knowledge, attitude, and skill, and secondly by sharing the results of the models with questionnaire items as indicators of the social tasks.

### Attitude

#### *Social Tasks as Indicators*

Comparing citizenship attitude across sex and SEP, metric invariance was reached. This justifies comparing associations between the construct citizenship attitude and its four underlying social tasks across sex and SEP. Comparing citizenship attitude across migration backgrounds, metric invariance could not be reached. Inspecting modification indices and standardised mean residuals pointed to no possibility of establishing a partial model. Therefore, only configural invariance was reached. This means that the path models specified for students with and without migration backgrounds are likely to follow the same pattern. For an overview, see Table 3.

**Table 3**

*Models with Social Tasks as Indicators of Citizenship Constructs*

| | | Configural | | Metric | | | Scalar | | | Partial Scalar | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | CFI | R | CFI | R | $\Delta \chi^2$ | CFI | R | $\Delta \chi^2$ | CFI | R | $\Delta \chi^2$ |
| Attitude | | | | | | | | | | | | |
| | Sex | + | + | + | + | + | - | - | - | | | |
| | SEP | + | + | + | + | + | + | + | - | + | + | + |
| | Migr. | + | + | + | + | - | | | | | | |
| Knowledge | | | | | | | | | | | | |
| | Sex | + | - | | | | | | | | | |
| | SEP | + | - | | | | | | | | | |
| | Migr. | + | - | | | | | | | | | |
| Skill | | | | | | | | | | | | |
| | Sex | + | + | + | + | + | + | + | - | + | + | + |
| | SEP | + | + | + | + | + | + | + | - | + | + | + |
| | Migr. | + | + | + | + | + | + | + | + | | | |

*Note.* Migr. = migration background; R = RMSEA. $\Delta \chi^2$ = Chi-square test of difference. A '+' for CFI indicates that the value for CFI is ≥ 0.95; a '-' indicates that the value is ≤ 0.95. A '+' for RMSEA indicates that the value for RMSEA is ≤ 0.08; a '-' indicates that the value is ≥ 0.08. A '+' for $\Delta \chi^2$ indicates that the p-value of the Chi-square test of difference is ≥ 0.05; a '-' for $\Delta \chi^2$ indicates that the p-value of the Chi-square test of difference is ≤ 0.05. The full tables can be found in Appendices 2.1, 2.2 and 2.3.

**Table 4**

*Models with Polytomous Questionnaire Items as Indicators of Social Tasks*

| | Configural | | Threshold | | | Metric | | | Partial Metric | | | Scalar | | | Partial Scalar | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CFI | R | CFI | R | $\Delta\chi^2$ | CFI | R | $\Delta\chi^2$ | CFI | R | $\Delta\chi^2$ | CFI | R | $\Delta\chi^2$ | CFI | R | $\Delta\chi^2$ |
| Attitude – Acting in a socially responsible manner | | | | | | | | | | | | | | | | | |
| Sex | + | + | + | + | + | + | + | + | | | | + | + | - | | | |
| SEP | + | + | + | + | + | + | + | - | + | + | + | | | | + | + | + |
| Migr. | + | + | + | + | + | + | + | + | | | | - | + | - | | | |
| Attitude – Dealing with conflicts | | | | | | | | | | | | | | | | | |
| Sex | + | + | + | + | + | + | + | - | | | | | | | | | |
| SEP | + | + | + | + | + | + | + | + | | | | + | + | + | | | |
| Migr. | + | + | + | + | + | + | + | + | | | | + | + | - | + | + | + |
| Attitude – Dealing with differences | | | | | | | | | | | | | | | | | |
| Sex | + | - | | | | | | | | | | | | | | | |
| SEP | + | - | | | | | | | | | | | | | | | |
| Migr. | + | - | | | | | | | | | | | | | | | |
| Skill – Acting in a socially responsible manner / Dealing with conflicts | | | | | | | | | | | | | | | | | |
| Sex | + | - | | | | | | | | | | | | | | | |
| SEP | + | - | | | | | | | | | | | | | | | |
| Migr. | + | - | | | | | | | | | | | | | | | |
| Skill – Dealing with differences | | | | | | | | | | | | | | | | | |
| Sex | + | + | + | + | + | + | + | + | | | | + | + | + | | | |
| SEP | + | + | + | + | + | + | + | + | | | | + | + | - | | | |
| Migr. | + | + | + | + | + | + | + | - | + | + | + | | | | + | + | - |

*Note.* Migr. = migration background; R = RMSEA. $\Delta\chi^2$ = Chi-square test of difference. A '+' for CFI indicates that the value for CFI is ≥ 0.95; a '-' indicates that the value is ≤ 0.95. A '+' for RMSEA indicates that the value for RMSEA is ≤ 0.08; a '-' indicates that the value is ≥ 0.08. A '+' for $\Delta\chi^2$ indicates that the p-value of the Chi-square test of difference is ≥ 0.05; a '-' for $\Delta\chi^2$ indicates that the p-value of the Chi-square test of difference is ≤ 0.05. The full tables can be found in Appendices 2.1, 2.2 and 2.3.

**Table 5**

*Models with Dichotomous Questionnaire Items as Indicators of Social Tasks*

| | Configural | | Scalar | | | Partial Scalar | | |
|---|---|---|---|---|---|---|---|---|
| | CFI | R | CFI | R | $\Delta \chi^2$ | CFI | R | $\Delta \chi^2$ |
| Knowledge – Acting democratically | | | | | | | | |
| Sex | + | + | + | + | + | | | |
| SEP | + | + | + | + | + | | | |
| Migr. | + | + | + | + | + | | | |
| Knowledge – Acting in a socially responsible manner | | | | | | | | |
| Sex | + | + | + | + | + | | | |
| SEP | + | + | + | + | + | | | |
| Migr. | + | + | + | + | + | | | |
| Knowledge – Dealing with conflicts | | | | | | | | |
| Sex | - | + | | | | | | |
| SEP | - | + | | | | | | |
| Migr. | + | + | + | + | - | + | + | + |
| Knowledge – Dealing with differences | | | | | | | | |
| Sex | + | + | - | + | - | + | + | + |
| SEP | + | + | + | + | + | | | |
| Migr. | + | + | + | + | + | | | |

*Note.* Migr. = migration background; R = RMSEA. $\Delta \chi^2$ = Chi-square test of difference. A '+' for CFI indicates that the value for CFI is $\geq 0.95$; a '-' indicates that the value is $\leq 0.95$. A '+' for RMSEA indicates that the value for RMSEA is $\leq 0.08$; a '-' indicates that the value is $\geq 0.08$. A '+' for $\Delta \chi^2$ indicates that the p-value of the Chi-square test of difference is $\geq 0.05$; a '-' for $\Delta \chi^2$ indicates that the p-value of the Chi-square test of difference is $\leq 0.05$. The full tables can be found in Appendices 2.1, 2.2 and 2.3.

***Items as Indicators***

Comparing the social tasks within citizenship attitude, (partial) metric invariance was reached in most cases. This allows for comparing the associations between the concerning social tasks and the underlying items across groups. Threshold invariance was achieved for *Attitude – Dealing with conflicts* across sex, but metric invariance could not be established. This means that the thresholds are the same across boys and girls, but the factor loadings between questionnaire items and the social task differ across boys and girls. For comparisons of *Attitude – Dealing with differences* across sex, SEP and migration background, the configural model indicated poor model fit. This suggests that we cannot assume that the path models specified for boys and girls, students with low and high SEP, and students with and without migration background, follow the same pattern; the same number of indicators relating to the same number of latent constructs. For an overview, see Table 4.

**Knowledge**

***Social Tasks as Indicators***

Comparing citizenship knowledge across sex, SEP and migration background, the configural model indicated poor model fit. This suggests that the data follow a different pattern in one of the groups. For example, it may be the case that one of the indicators of citizenship knowledge, e.g., dealing with differences, is an indicator of citizenship knowledge for students with a low SEP but not for students with a high SEP. For an overview, see Table 3.

***Items as Indicators***

Comparing the social tasks within citizenship knowledge, the minimum of (partial) scalar invariance was reached in most cases. This means that we can meaningfully compare the mean scores of the concerning social tasks across groups. For comparisons of the social task *Knowledge – Dealing with conflicts* across sex and SEP, the configural model indicated poor model fit. For an overview, see Table 5.

**Skill**

**Social Tasks as Indicators**

Comparing citizenship skill across sex and SEP, metric invariance was reached. This justifies the comparison of citizenship skill and its associations with the underlying social tasks across boys and girls and students with a low and high SEP. Comparing citizenship skill across migration backgrounds, scalar invariance was reached. This means that we found no evidence that the intercepts differ across students with and without migration backgrounds, and we can meaningfully compare the mean scores of citizenship skill. For an overview, see Table 3.

**Items as Indicators**

Comparing the social tasks within citizenship skill, the required minimum of (partial) metric invariance was reached for comparisons of *Skill – Dealing with differences* across sex, SEP, and migration background. This means we may compare the associations between the social task and its underlying questionnaire items across groups. However, for comparisons of *Skill – Acting in a socially responsible manner/Dealing with conflicts* across sex, SEP, and migration background, the configural model indicated poor model fit. This means that the path models will likely follow a different pattern across the groups. For an overview, see Table 4.

**Permutation Tests**

The permutation tests resulted predominantly in non-significant p-values, indicating that the significant Chi-square tests of overall model fit depended on the same underlying reasons across groups (also see Appendix 2.5). However, in the comparison of *Skill – Dealing with differences* across sex and *Attitude – Acting in a socially desirable way* across migration background, the p-value was respectively 0.045 and 0.012. Here, the significant Chi-square tests of overall model fit may depend on different underlying reasons across groups. This means that we need to be precautious when interpreting the results of these two comparisons.

**Exploratory Analyses**

We established a partial metric model for *Skill – Dealing with differences* across migration backgrounds. For *Attitude – Acting in a socially responsible way* across SEP, *Knowledge – Dealing with conflicts* across migration background, and *Knowledge – Dealing with differences* across sex, we established partial scalar models. This means that in these models, some indicators were less invariant than others and hindered reaching a higher level of measurement invariance. We performed exploratory analyses to zoom in on these indicators by looking at the content of the questionnaire items, the response frequencies, the factor loadings and the item-rest correlations (also see Appendix 2.4). In most cases, we found that the difference in factor loadings across groups was the largest for the less invariant indicators.

For example, comparing *Attitude – Acting in a socially desirable way* across SEP, we established a partial scalar model with freed equality constraints on the second and fourth indicator intercepts. Exploratory analyses of these two indicators demonstrated that the factor loading was higher for students with a high SEP for the second indicator. For the fourth indicator, the opposite was true. This means that the second indicator ('*If a classmate is being called names in the streets, I want to stick up for him or her'*) more strongly relates to the construct *Attitude – Acting in a socially desirable way* for students with a high SEP as compared to students with a low SEP. Contrary, the fourth indicator ('*You should say sorry if you did*

*something that hurt another person*') more strongly relates to the construct for students with a low SEP as compared to students with a high SEP.

We also found a difference in factor loadings across groups. For example, comparing *Efficacy – Dealing with differences* across migration background, we established a partial metric model with freed equality constraints on the factor loading of the first indicator. Exploratory analyses demonstrated that for the first indicator (*'How good are you at… adapting your behaviour to the rules and habits of others?'*), the factor loading was larger for students without a migration background, whereas for the remaining indicators (*'How good are you at… behaving normally in an unfamiliar environment?'; '... adapting your language to the person you are speaking with?'; '... considering the wishes of others when making a decision together?')*, the factor loadings were larger for students with a migration background.

The permutation tests resulted predominantly in non-significant p-values, indicating that the significant Chi-square tests of overall model fit depended on the same underlying reasons across groups. However, in the comparison of *Skill – Dealing with differences* across sex and *Attitude – Acting in a socially desirable way* across migration background, the p-value was respectively 0.045 and 0.012. Here, the significant Chi-square tests of overall model fit may depend on different underlying reasons across groups. This means that we need to be precautious when interpreting the results of these two comparisons.

## Discussion

Standardised questionnaires are widely used to measure the outcomes of citizenship education in terms of knowledge, attitude, and skill (Ireland et al., 2006; Schulz et al., 2018). The outcomes of these questionnaires are often compared across groups based on student characteristics, such as boys and girls. Insight into the outcomes of citizenship education is important to evaluate the effectiveness of the content and delivery methods of the curriculum. Moreover, insight into the extent to which these outcomes differ across groups helps to ensure that all students have equal opportunity to acquire citizenship competences. However, a prerequisite for valid cross-group comparisons based on standardised questionnaires entails addressing whether the instrument measures the same across groups and later points in time. This can be done by examining the measurement invariance (Meredith, 1993).

The establishment of measurement invariance is, however, not a static given because the interpretation of constructs can change over time. This may particularly be the case for value-sensitive concepts like citizenship and, accordingly, citizenship education, which is considered a dynamic construct that takes on meaning in context and moves along with changes in society. Hence, it is important to periodically examine the measurement invariance of

measurement instruments capturing citizenship competences. In line with this rationale, this study aimed to investigate the measurement invariance of the CCQ (Geijsel et al., 2012; Ten Dam et al., 2011), which is a standardised questionnaire based on a generic conceptualisation of citizenship competences used over a long period. Schools use the instrument to gain insight into students' citizenship knowledge, attitudes, and skills across four social tasks: acting democratically, acting in a socially responsible manner, dealing with conflicts and dealing with differences. To check whether this instrument over time still measures the same across groups in recent samples, we conducted comparisons within three student characteristics which are known to relate to robust differences in citizenship competences: sex, SEP, and migration background (Dijkstra et al., 2015; Geijsel et al., 2012).

Our study showed that in two-thirds of the comparisons across groups, the meaningful minimum of at least (partial) metric invariance (Little, 2013) was reached. This allows for comparisons of the associations between latent constructs and their indicators across groups (i.e., 'the relation between indicator 1 and latent construct X is stronger for boys than for girls'). We even reached (partial) scalar invariance across groups in some comparisons. This allows for comparing latent means across groups (i.e., 'girls, on average, obtain a higher score on latent construct X than boys'). Our findings also mark a warning for researchers conducting multi-group comparative analyses based on citizenship constructs or social tasks where the minimum required level of measurement invariance could not be established. In this study, this applied to one-third of the comparisons, being one citizenship construct (i.e., knowledge) and two social tasks (i.e., Skill – Acting in a socially responsible manner/Dealing with conflicts, and Attitude – Dealing with differences), where fit indices of the configural model indicated overall poor model fit. This means that, for these three aspects, we have no grounds to assume that the instrument measures the same across sex, SEP and migration background, and we need to be cautious when making cross-group comparisons. Nevertheless, the results on these aspects can still be used to describe all students in a class or school as a whole. Moreover, it is advised to assess measurement invariance over these aspects again when using new samples. Whereas we advocate periodic assessment of measurement invariance of *all* citizenship constructs, our findings underscore this is indeed important for future studies examining the construct and social tasks mentioned above.

In some comparisons, specific indicators hindered establishing a higher level of invariance. Therefore, we searched for partial invariance where these particular indicators remained on the lower level of invariance while the other indicators were specified at a higher level of invariance. Exploratory analyses of the partial models pointed to no unidirectional

reason why these particular indicators may have been answered differently across groups. That is, we could not identify measurement non-invariance due to either of the three explanations mentioned by Van de Vijver (2013) and Isac et al. (2019): (1) some group members not considering some underlying indicators to be indicative of the construct; (2) cultural or linguistic differences across group members; or (3) a specific response-style of some group members. To improve the questionnaire regardless, we suggest performing additional qualitative data analysis on these non-invariant indicators to see whether this does reveal unidirectional leads of how to improve the indicators. This can be done in individual interviews or panel group interviews with students originating from all comparison groups. Alternately, new indicators can be added to the scales and tested among students to see whether they do show measurement invariance across groups. Eventually, new invariant indicators can replace the non-invariant indicators.



While the number of studies that assessed the measurement invariance of citizenship competences within a country is scarce, we can, to some extent, compare our findings to the studies that performed a between-country assessment of measurement invariance, which is more common in large-scale international assessments of citizenship competences. An example is ICCS (Schulz et al., 2016), where most scales were invariant across countries. The technical report places caution in comparing only two scales (i.e., students' perceptions of the importance of citizenship behaviours; and students' attitudes toward civic institutions and their country of residence). Isac et al. (2019) also used data from ICCS and, similarly, found that most scales measuring students' attitudes towards immigrants were invariant across countries. Only the scale measuring students' attitudes towards equal rights for immigrants was more heterogeneous across countries. The indicator 'Immigrants should have the opportunity to continue their own customs and lifestyle' was the weakest in the scale. In addition, Munck et al. (2018) found that the measurement of students' attitudes toward immigrants was largely invariant over time (from 1999 to 2009) across countries. However, the indicator 'immigrants should have the opportunity to keep their own language' was invariant in only 36 out of 92 countries. Whereas the studies of Isac et al. (2019) and Munck et al. (2018) showed that between-country measurement invariance was more difficult to establish in measuring students' attitude in dealing with differences, our findings of non-invariance in Attitude – Dealing with differences indicate the same in within-country comparisons. However, these previous studies do not comply with the non-invariance that we found for the construct Knowledge and the social task Skill – Acting in a socially responsible manner/Dealing with conflicts.

Although we grounded our findings on extensive analyses, it is important to bear in mind the limitations of this study. Caution needs to be placed on the way we composed the groups. We cared to comply with scholars who underscored that the robustness and precision of measurement invariance analyses increase as sample sizes increase (Koh & Zumbo, 2008; Meade, 2005; Meade & Lautenschlager, 2004). Therefore, we combined data from students of five subsequent years sharing the same student characteristics, and consequently, we were limited in comparing measurement invariance over time. However, the comparison over time is an important lead for future research. We were also somewhat limited in the extent to which we could differentiate within each group. For example, we combined data from students whose parents were born in countries other than the Netherlands into 'having a migration background'. Whereas it is not inconceivable that the effect on measurement invariance differs within these countries, this operationalisation suffices to provide insight into the main effect of migration background in assessing measurement invariance. The same may hold for SEP, where we combined different attained parental educational levels into low SEP.

Despite the limitations, this study contributes to the empirical knowledge base of the assessment of measurement invariance in instruments for citizenship competences. Whereas previous research on the assessment of measurement invariance in instruments for citizenship competences focused on between-country comparisons (Isac et al., 2019; Schulz et al., 2016), the assessment of measurement invariance is equally important in within-country comparisons (Steinmetz et al., 2009). This study meets this importance. However, at the same time, the findings of this study make clear that the results of the assessment of measurement invariance are not static. Citizenship and, as a result, citizenship education is dynamic and subject to changes in society and schools. These changes may lead to changes in how an instrument aligns with its context and underscores the importance of periodic assessment of measurement invariance.

# Chapter 3

# Understanding Effective Citizenship Education in Special Needs Education: Comparing Regular Primary Education with Special Needs Primary Education




## Abstract

Research conducted in regular primary education (RPE) identified aspects of effective citizenship education, such as the teacher-student relationship and an open classroom climate for discussion. In schools for special needs primary education (SNPE), such insight is missing.

This study examined differences in citizenship competences between students attending SNPE versus RPE and the moderating role of several characteristics of citizenship education. To do so, we conducted multilevel analyses based on 3782 students in 179 schools.

Students in SNPE obtained lower scores on civic knowledge, self-estimated skills, and attitude. Whereas the role of some characteristics was not as expected and ask for further research, this study shed light on several important factors, such as the attention for specific learning goals in citizenship education, that form an important basis for further research into characteristics of effective citizenship education in SNPE.


*Keywords:* Citizenship education; Citizenship competences; School effectiveness; Special needs education; Primary education.

*This chapter is based on: Hoek, L., Slijkhuis, E. G. J., Munniksma, A., & Dijkstra, A. B. (2023). Understanding effective citizenship education in special needs education: Comparing regular primary education with special needs primary education. *Accepted for publication in Citizenship Teaching & Learning.*

## Introduction

Many countries have formalised the importance of strengthening social cohesion via citizenship education in schools in policy, national curricula or legislation (Eurydice, 2017). Citizenship education is an educational domain that prepares students to participate in a democratic and pluriform society (Ten Dam et al., 2011), for example, by promoting fundamental democratic values and fostering mutual respect. The learning outcomes of citizenship education are often referred to as citizenship competences and are usually operationalised as knowledge, skills, and attitude (e.g. Schulz et al., 2016; Ten Dam et al., 2011). Following the increased attention to citizenship education in policy and educational practice, a considerable amount of research has focused on the learning outcomes of citizenship education (Geboers et al., 2013) and on factors that influence the effectiveness of citizenship education (Coopmans et al., 2020; Dijkstra et al., 2015). It is, for example, often asserted that various aspects of the school climate (Sampermans et al., 2018; Thapa et al., 2013; Vieno et al., 2005) contribute to the effectiveness of citizenship education. These aspects include a classroom climate that fosters dialogue and critical debate on political and societal issues (Geboers et al., 2013; Knowles et al., 2018) and inclusive and mutually respectful teacher-student and student-student relationships (Isac et al., 2014; Wanders, Van der Veen, et al., 2020; Wanders, Dijkstra, et al., 2020).

The increased scientific attention to the effective aspects and outcomes of citizenship education, however, is predominantly focused on schools for regular education. In special needs primary education (SNPE), research into the effectiveness of citizenship education is largely missing. Only recently, a descriptive study among Dutch students in SNPE demonstrated that students in SNPE generally obtained lower scores on citizenship competences than students in regular primary education (RPE) (Slijkhuis et al., 2023). This finding was an important starting point for the present study, as more insight into the outcomes and characteristics of citizenship education in SNPE helps tailoring the contents of citizenship education to what students in SNPE need. For example, research in regular education demonstrates that positive teacher-student and student-student relationships promote students' citizenship competences (Sampermans et al., 2018; Wanders, Dijkstra et al., 2020; Wanders, Van der Veen et al., 2020). At the same time, students with behavioural difficulties are more at risk of developing negative relationships with the teacher (Roorda et al., 2021) and peers (Pinto et al., 2019). As such, it is unclear whether the fact that students in SNPE obtained lower scores on citizenship competences (cf. Slijkhuis et al., 2023) is influenced by characteristics of

citizenship education, such as whether students perceive positive relationships with the teacher and peers.

Hence, the aim of the present study is twofold. First, the aim is to examine whether characteristics of citizenship education are different in SNPE versus RPE and to give insight into the relationship between attending SNPE and citizenship competences. The second aim is to investigate whether the relationship between attending SNPE and citizenship competences in terms of knowledge, skills, and attitude changes under the influence of the characteristics of citizenship education in SNPE and RPE. Following these aims, the insights of this study provide a basis for developing and improving SNPE citizenship education, based on empirical knowledge resulting from research conducted among their specific student population. We do so by investigating the following research questions:

1. To what extent are characteristics of citizenship education different in SNPE versus RPE?
2. To what extent is attending SNPE associated with citizenship competences in terms of knowledge, skills, and attitude?
3. To what extent do characteristics of citizenship education moderate the relationship between attending SNPE and citizenship competences in terms of knowledge, skills, and attitude?

This study used data from two related studies conducted in SNPE and RPE in the Netherlands. In both studies, skills are operationalised in two components: applied skills and self-estimated skills. Primary analyses using these data already demonstrated that the relationship between attending SNPE and citizenship competences is generally negative (cf. Slijkhuis et al., 2023), but, to provide a complete picture, the second research question was also included in the present study.

Like in many other European countries (Schwab, 2021), policy in the Netherlands is steered towards the inclusion of students with special educational needs (SEN) in regular education. Still, it explicitly promotes finding education that fits students' capacities regardless of whether this is in regular education (Van der Veen et al., 2010). In the Netherlands, students with SEN may attend RPE, SNPE or special primary education when they are between 4 and 12 years old (Van der Veen et al., 2010). Students without learning or behavioural difficulties usually attend RPE; students with moderate learning or behavioural difficulties usually attend special primary education; students with more severe learning or behavioural difficulties usually participate in

SNPE[5]. This study is focused on a comparison between, on the one hand, students attending RPE and, on the other hand, students with moderate to severe learning or behavioural difficulties attending either special primary education or SNPE.

## Theoretical Framework

This section starts with characterising students who attend SNPE by elaborating on their experienced challenges in educational settings. Next, based on empirical knowledge of 'what works' in citizenship education among students in RPE and research on students' learning and behavioural difficulties in SNPE, we discerned four characteristics of effective citizenship education that could alter the relationship between attending SNPE and citizenship competences. They are: experienced unsafety in the school environment, the teacher-student and student-student relationship, and an open classroom climate for discussion. We close each paragraph by providing a hypothesis on the moderating role of these characteristics in the relationship between attending SNPE and citizenship competences. Previous research indicated that more variance could be attributed to the school in knowledge and applied skills as opposed to self-estimated skills and attitude (De Schaepmeester et al., 2022b; Isac et al., 2011; Janmaat, 2011). Hence, in these hypotheses, the moderation effects are expected to be stronger for knowledge and applied skills than self-estimated skills and attitude.

### Special Educational Needs

Whereas there seems to be no clear and universal definition of SEN among students (Schwab, 2021), it generally refers to 'pupils with various (combinations of) impairments and/or difficulties in participating in education' (Pijl et al., 2008, p. 389). These impairments or difficulties include physical and mental impairments and learning and behavioural difficulties. The present study pertains to the latter two. Examples of learning and behavioural difficulties include, but are not limited to: externalising problem behaviour, internalising problem behaviour, a problematic attitude to work (i.e., lack of motivation, effort and concentration; slow pace of work), having a speech, language or numeracy disorder (i.e., speech or language disorder, dyslexia, dyscalculia), being gifted, autism spectrum disorder, or being behind in literacy and or numeracy (Van der Veen et al., 2010). Some scholars pose that citizenship competences among students with SEN are best learnt through experience and interactions with others (Alevriadou & Lang, 2011). According to others, active citizenship of

[5] In the Netherlands, special education is divided in four domains: (1) for blind and visually impaired students; (2) for deaf and hearing impaired students and students with a language development disorder; (3) for students who are physically and mentally impaired and students who are long-term sick; and (4) for students with severe learning and behavioural difficulties. This study did not include students in SNPE who are blind or visually impaired, deaf or hearing impaired, physically and mentally impaired, and long-term sick students.

students with SEN is ideally oriented towards abilities rather than disabilities and should focus on autonomy, self-awareness and self-direction (Reiter & Schalock, 2008). This prudently points to the assumption that citizenship education in SNPE is focused on the intrapersonal and interpersonal aspects of citizenship education and less towards the societal aspect (cf. Ten Dam & Volman, 2007).



**Unsafety in the School Environment**

Unsafety in the school environment is commonly understood as the presence of unsafe social situations in the school environment and is seen as an aspect of the school climate (cf. Thapa et al., 2013). Perceptions of the school environment (measured by safety and inclusion in the school, bullying experiences, and clear expectations for behaviour) were significantly related to students' academic success (Gietz & McIntosh, 2014) and promoted student learning (Devine & Cohen, 2007). In addition, the school climate was linked to students' future voting intentions (Sampermans et al., 2018), which is considered an important part of citizenship competence.

Students who attend SNPE, particularly those experiencing behavioural difficulties, generally have less developed socio-emotional competences than their peers who attend RPE (Frostad & Pijl, 2007). In addition, students with internalising and externalising behavioural difficulties (Farmer et al., 2015; Ledwell & King, 2015; Rose & Espelage, 2012) and learning disabilities (Mishna, 2003) are at risk of involvement in bullying behaviour. The presence of bullying behaviour is an important indicator of unsafety in the school environment (Gietz & McIntosh, 2014).

In line with research demonstrating that (a) perception of safety in the school environment promotes learning outcomes in general (Devine & Cohen, 2007; Gietz & McIntosh, 2014) and in the context of citizenship education (Sampermans et al., 2018), and (b) research demonstrating that students who have learning or behavioural difficulties are more at risk to experience unsafety in the school environment (Farmer et al., 2015; Frostad & Pijl, 2007; Ledwell & King, 2015; Mishna, 2003; Rose & Espelage, 2012), it is hypothesised that the relationship between attending SNPE and citizenship competences is more negative for students who experience more unsafety in the school environment (Hypothesis 1).

**Teacher-student and Student-student Relationship**

Interpersonal relationships between students and teachers and among students, also referred to as the teacher-student relationship and the student-student relationship, have been extensively studied. There is consistent and robust empirical evidence demonstrating that affective teacher-student relationships are related to students' behavioural, emotional and

academic development (Hamre & Pianta, 2001) and school engagement (Roorda et al., 2021). In addition, various scholars have demonstrated the relationship between the quality of student-student relationships and academic achievement (Wentzel, 2017) and feelings of safety and motivation to participate in the classroom (Baker et al., 2008; Hamre & Pianta, 2001). The teacher-student and student-student relationships were also studied in the context of citizenship education, where positive relationships promoted students' intention to go voting in the future (Sampermans et al., 2018) and were related to students' societal involvement (Wanders, Dijkstra, et al., 2020; Wanders, Van der Veen, et al., 2020).

Students with internalising behaviours (e.g., shyness, social withdrawal, anxiety, and depression) and externalising behaviours (e.g., overactivity, impulsiveness, and aggression) are considered at risk for experiencing difficulties in forming affective relationships with their teachers (Baker et al., 2008; Freire et al., 2020; Zee & Roorda, 2018) and with students (Carter et al., 2014; Pinto et al., 2019; Schwab, 2019). For example, in a study examining differences in teacher-student relationship quality and school engagement between boys from RPE and boys attending special education for students with autism spectrum disorders, the authors found that the boys from special education reported more conflict in the teacher-student relationship and less school engagement as compared to the boys in RPE (Roorda et al., 2021). In addition, in a study comparing peer relations of students who attend SNPE versus RPE, students attending SNPE had lower levels of peer acceptance, fewer reciprocated friendships and were less integrated into peer groups (Pinto et al., 2019). This was also confirmed in a literature review, where students with special needs in secondary education were often found to have fewer social interactions with their peers, to be less popular than their classmates and to be more often neglected or rejected (Bossaert et al., 2013).

In line with research demonstrating that (a) positive teacher-student and student-student relationships promote learning outcomes in general (Baker et al., 2008; Hamre & Pianta, 2001; Roorda et al., 2021; Wentzel, 2017) and in the context of citizenship education (Sampermans et al., 2018; Wanders, Dijkstra, et al., 2020; Wanders, Van der Veen, et al., 2020), and (b) research demonstrating that students with learning and behavioural difficulties are more at risk for experiencing negative relationships with the teacher and fellow students (Freire et al., 2020; Pinto et al., 2019; Zee & Roorda, 2018), it is hypothesised that the relationship between attending SNPE and citizenship competences is more positive for students who experience more positive teacher-student (Hypothesis 2) and student-student (Hypothesis 3) relationships.

**Open Classroom Climate for Discussion**

An open classroom climate can be defined as a classroom climate that 'fosters discussion of controversial issues, encourages students to present diverse opinions, and emphasises respect of those opinions from teachers and students' (Godfrey & Grayman, 2014, p. 1802). It is seen as a requirement for citizenship education to be effective (Knowles et al., 2018; Maurissen, 2018), and it is one of the most constant factors positively influencing students' citizenship competences (Geboers et al., 2013; Maurissen, 2018). Previous research, for example, shows a positive effect on citizenship knowledge (Alivernini & Manganelli, 2011; Campbell, 2008; Martens & Gainous, 2013; Maurissen et al., 2018; Persson, 2015), citizenship skills (Claes et al., 2017; Kuang et al., 2018; Manganelli et al., 2015), and citizenship attitude (Coopmans et al., 2020).

To our knowledge, no empirical research is available on the open classroom climate for discussion fostering citizenship competences in SNPE. However, it is not inconceivable that some of the skills necessary for an open classroom climate (e.g., being able to listen to others, being able to present one's own opinion on a subject matter, being able to 'agree to disagree', and so on) may be extra challenging for students facing learning or behavioural difficulties. In that respect, internalising behaviour such as shyness, anxiety, lack of self-confidence or fear of failure may make it hard for students to present their own opinion, whereas externalising behaviour such as having a disruptive effect on others or aggressive behaviour towards classmates may make it hard for students to experience the classroom climate as sufficiently open and safe for having a discussion (Van der Veen et al., 2010).

In line with research demonstrating that (a) an open classroom climate fosters the acquisition of citizenship knowledge (Alivernini & Manganelli, 2011; Campbell, 2008; Martens & Gainous, 2013; Maurissen et al., 2018; Persson, 2015), skills (Claes et al., 2017; Kuang et al., 2018; Manganelli et al., 2015), and attitude (Coopmans et al., 2020) and (b) research characterising the behavioural and learning difficulties students in SNPE may experience (Van der Veen et al., 2010), possibly also in establishing and contributing to an open classroom climate for discussion, it is hypothesised that the relationship between attending SNPE and citizenship competences is more positive for students who perceive the classroom climate to be more open for discussion (Hypothesis 4).

**Other Characteristics of Effective Citizenship Education in Regular Education**

We also included nine additional characteristics of citizenship education in our analyses based on their demonstrated effectiveness in studies among students attending (predominantly secondary) regular education, but not necessarily on research on the learning and behavioural

difficulties of students in SNPE. Therefore, we chose to omit hypotheses here, and this study can be considered exploratory concerning these additional characteristics.

First, more attention paid to citizenship education in school staff meetings was positively related to citizenship attitudes and skills of students (Coopmans et al., 2020). The authors also found that more agreement among teachers on the delivered content of citizenship education was positively related to citizenship attitudes and that monitoring the development of citizenship competences was negatively (Coopmans et al., 2020) and positively (Keating et al., 2010) associated with citizenship knowledge, skills, and attitude. The effectiveness of citizenship education was also linked to the ability of schools to develop policies and take concrete actions to improve their teaching practice and learning environment (Kyriakides et al., 2010). In addition, having a clear vision of citizenship education as a school positively influenced students' citizenship skills (Dijkstra et al., 2015).

## Methodology

### Participants

We used data from two separately conducted but content-related studies initiated by the Inspectorate of Education in the Netherlands in spring 2020 (RPE study) and spring 2021 (SNPE study). In both cases, the main sample was drawn from the total population of schools. Students in selected schools were excluded from participation if they were diagnosed with a language development disorder, if they resided in the Netherlands for less than two years at the time of the study and could not sufficiently master the Dutch language, and if they had an estimated IQ below 75 (the latter only in the RPE study). Eventually, in the RPE study, we used data from 2237 students (52% girls, mean age is 11,54) and 117 teachers (70% women) from 94 schools. In the SNPE study, we used data from 1545 students (28% girl, mean age is 11,94) and 162 teachers (77% woman) from 85 schools. Based on the total number of participating students, teachers and schools, as well as the considered criteria such as school region and estimated future educational paths of students, both samples can be regarded as representative of the SNPE and RPE populations. The descriptive statistics of the sample can be found in Appendix 3.1.

### Procedure

After schools agreed upon participation, parents received an information letter and were allowed to deny their children's involvement. Before students filled in the questionnaire, they were ensured anonymity and allowed to quit participation at all times. After the instructions were given, students took approximately one hour to complete the questionnaire.

For students in RPE, a set of 208 items was divided into eight questionnaires, with each

52 items asking about students' citizenship knowledge, applied skills, self-estimated skills, and attitude. Each student received one out of eight questionnaires. The incomplete research design was chosen because it enabled receiving results on a larger set of items than a complete design could provide. Scaling was based on item calibration using Rasch models. The technical report extensively describes this procedure (Slijkhuis et al., 2021). The SNPE study used a complete research design because the item load per student was more limited here, and the risk of insufficient observations per item in an incomplete research design was estimated to be too big due to the COVID-19 pandemic. Instead, students in SNPE received a subset of 60 carefully selected items based on their coverage of the intended topics and competences.



**Measures**

Knowledge, applied skills, self-estimated skills and attitude are derived from the conceptual framework of Ten Dam et al. (2011). The dependent variables cover the so-called social tasks 'acting democratically', 'acting in a socially responsible way', 'dealing with conflicts', and 'dealing with differences'. An overview of the dependent and independent variables is provided in Table 1. Following previous research showing that sex, age, ethnical background and socioeconomic position are related to differences in citizenship competences (Geboers et al., 2013; Hoskins et al., 2017; Isac et al., 2014), we used the following control variables: sex, age, language spoken at home (not (only) Dutch or only Dutch) and advised educational level (education for students with severe learning difficulty; work/daytime activities; practical job-oriented education; practical pre-vocational secondary education; theoretical pre-vocational secondary education; senior general secondary education; pre-university secondary education). In addition, we controlled for IQ (below 75 or not) because, for 23% of the students in the SNPE study, the IQ was below 75.

**Analytic Plan**

We performed univariate multilevel modelling using the lme4 package (Bates et al., 2015) in R version 4.0.5 (R Core Team, 2021). By doing so, we could control for the fact that students were nested within schools. We first constructed an empty model with attending SNPE as the predictor of citizenship knowledge, applied skills, self-estimated skills and attitude. Next, we added the control variables to the model. In the third model, we added unsafety in the school environment, the teacher-student and student-student relationship, and students' perceived openness of classroom climate as predictors. In the fourth model, we separately added the interaction effects between these predictors and attending SNPE. In the fifth model, we built upon the second model by adding the following predictors: attention to personal development and social competence, attention to promoting social involvement and

participation, attention to the promotion of dealing with differences and fostering of an open attitude towards differences, open classroom climate (teachers' perception), monitoring of learning outcomes, formulation of a vision, formulation of learning goals, agreement on the curriculum of citizenship education, and the importance of citizenship education. In the last model, we separately added the interaction effects between each predictor and attending SNPE.

**Data Preparation**

We observed 5.36% missing observations that showed no clear pattern of missingness. We performed multivariate imputation by chained equations on our missing observations using the mice package in R (Van Buuren & Groothuis-Oudshoorn, 2011).

## Results

### Preliminary Analyses

The correlation coefficients of the relevant variables are displayed in Appendix 3.2. To check for multicollinearity, we assessed the variation inflation factors (VIFs) (Mansfield & Helms, 1982). All independent variables had VIFs between 1 and 2. We also constructed null models for each outcome variable to inspect the intraclass correlations (ICCs). We found ICCs of respectively 0.18, 0.13, 0.08 and 0.10. This means that 18% of the variance between students' scores in knowledge can be attributed to the school. Likewise, 13% of students' applied skills, 8% of their self-estimated skills and 10% of their attitude can be attributed to their school. Via a subset of the data, we looked at the difference in ICCs between the students attending SNPE and RPE. The results show 4% (RPE) versus 15% (SNPE) of knowledge, 3% (RPE) versus 13% (SNPE) of applied skills, 6% (RPE) versus 3% (SNPE) of estimated skills, and 8% (RPE) versus 7% (SNPE) of attitude can be attributed to the school. This shows that the school attributes more to estimated skills and attitude of students attending RPE. In contrast, for students attending SNPE, the school attributes more to their knowledge and applied skills.

**Table 1**

*Overview of Dependent and Independent Variables*

| Construct | N items | Answer options | Cronbach's α | Example item |
|---|---|---|---|---|
| Knowledge | 15 (SNPE) or a set of 20 out of 78 items (RPE) | Multiple choice with one correct answer | α = 0.62 in SNPE | Which of the following is true? (a) During Ramadan, Muslims are allowed to eat extra food; (b) The cow is a sacred animal for Hinduists in India; (c) Catholic priests are allowed to marry their partners. |
| Applied skills | 10 (SNPE) or a set of 9-13 out of 44 items (RPE) | Multiple choice with one correct answer | α = 0.60 in SNPE | A group of children want a playground in their street. How can they achieve this by using a democratic procedure? (a) Gather signatures and hand these over to the municipality; (b) Find a place themselves and play over there; (c) Have a seat in the mayor's garden and wait until he or she concedes. |
| Self-estimated skills | 14 (SNPE) or a set of 2-8 out of 20 items (RPE) | Likert scale ranging from 1 (not good at all) to 4 (very good) | α = 0.74 in SNPE | How good are you at… respecting the opinion of others? |
| Attitude | 21 (SNPE) or a set of 15-17 out of 64 items (RPE) | Likert scale ranging from 1 (does not apply at all to me) to 4 (applies completely to me) | α = 0.86 in SNPE | If I have a picnic with my friends at the park, it is normal that we clean up after ourselves. |
| Unsafety in the school environment[a] | 6 items | Likert scale ranging from 1 (never) to 4 (often) | α = 0.83 in SNPE; α = 0.73 in RPE | A student called you names. |
| Teacher-student relationship[b] | 7 items | Likert scale ranging from 1 (completely incorrect) to 5 (completely correct) | α = 0.87 in SNPE; α = 0.87 in RPE | The teacher usually knows how I feel. |
| Student-student relationship[b] | 6 items | Likert scale ranging from 1 (completely incorrect) to 5 (completely correct) | α = 0.83 in SNPE; α = 0.73 in RPE | I would rather be in another class. |
| Open classroom climate for discussion (student perception)[a] | 6 items | Likert scale ranging from 1 (never) to 4 (often) | α = 0.80 in SNPE; α = 0.73 in RPE | The teacher helps the students in our class to form their own opinion. |
| Attention to personal development and social competence | 9 items | Likert scale ranging from 1 (not at all) to 5 (a lot) | α = 0.86 in SNPE; α = 0.88 in RPE | To what extent do you pay attention to <<learning to cooperate>> in citizenship education? |
| Attention to promoting social involvement and participation | 8 items | Likert scale ranging from 1 (not at all) to 5 (a lot) | α = 0.86 in SNPE; α = 0.70 in RPE | To what extent do you pay attention to <<taking different perspectives of a matter>> in citizenship education? |
| Attention to the promotion of dealing with differences and | 6 items | Likert scale ranging from 1 (not at all) to 5 (a lot) | α = 0.86 in SNPE; α = 0.78 in RPE | To what extent do you pay attention to <<countering of discrimination and prejudice>> in citizenship education? |

| | | | | |
|---|---|---|---|---|
| fostering an open attitude towards differences | | | | |
| Open classroom climate for discussion (teacher perception)[1] | 6 items | Likert scale ranging from 1 (almost none) to 4 (almost all) | α = 0.71 in SNPE; α = 0.72 in RPE | How many students in grade 6 propose topics to discuss in class? |
| Monitoring of learning outcomes of citizenship education | 5 items | 1 (almost) never), 2 (sometimes) and 3 (often) | NA | To what extent do you monitor outcomes in the domain of << attitudes (for example, sense of responsibility, taking others into account)>>? |
| Formulation of a vision in citizenship education | 1 item | 1 (everyone does this on their own, 2 (there are global ideas but no concretised vision), 3 (there is a shared vision, but it is not documented) and 4 (there is a documented vision) | NA | To what extent is citizenship education concretised into a vision? |
| Learning goals for citizenship education | 5 items | 1 (yes) or 2 (no) | α = 0.77 in SNPE; α = 0.75 in RPE | In our school, we have concrete agreement on the planning and activities in citizenship education. |
| Agreement on the curriculum of citizenship education | 9 items | Likert scale ranging from 1 (do not agree at all) to 5 (completely agree) | α = 0.76 in SNPE; α = 0.87 in RPE | In our school, the school board has insight into how teachers implement citizenship education. |
| Importance of citizenship education | 2 items | Likert scale ranging from 1 (totally disagree) to 5 (totally agree) | NA | I find citizenship education to be a task for the school. |

*Note.* The subscript 'a' refers to scales derived from Schulz et al. (2016); 'b' refers to scales derived from Peetsma et al. (2001). The RPE study used eight booklets with 52 items each. Each item was asked in two different booklets, leaving a response of approximately 484 observations per item (not considering missing observations). Hence, we opted not to provide values for Cronbach's α for the outcome variables of the RPE study.

**Main Analyses**

***Research Question 1: The Difference in Characteristics of Citizenship Education between Students in RPE versus SNPE***

Two sample T-tests were conducted to compare the characteristics of citizenship education between students in SNPE versus RPE. The results in Table 2 display that, whereas students in SNPE experienced the classroom climate to be more open to discussion, they also experienced more unsafety in the school environment and perceived lower teacher-student and student-student relationships as compared to students attending RPE. Contrary to students' perception, teachers in SNPE experienced the classroom climate to be less open to discussion than teachers in RPE. We also observed lower mean scores in SNPE for teachers' attention to promoting social involvement and participation in citizenship education (albeit a very small difference in means), for monitoring learning outcomes, for the formulation of a vision, for the formulation of concrete learning goals, and the agreement on the curriculum throughout different educational levels. We observed higher mean scores in SNPE for teachers' attention to personal development and social competence in citizenship education, teachers' promotion of dealing with differences and fostering an open attitude towards differences in citizenship education (albeit a very small difference in means), and the importance of citizenship education (also a very small difference in means).

***Research Question 2: The Difference in Citizenship Competences between Students in RPE versus SNPE***

Controlling for sex, age, language spoken at home, advised educational level and having an IQ below 75, attending SNPE was negatively related to knowledge ($b$ = -0.04, $p$ = 0.03), self-estimated skills ($b$ = -0.09, $p < 0.01$), and attitude ($b$ = -0.06, $p$ = 0.01). For applied skills, it did not matter whether students attended SNPE or RPE. The results of these models are provided in Table 3.

**Table 2**

*Two Sample T-tests Dependent Variables*

| | RPE<br>*M(sd)* | SNPE<br>*M(sd)* |
|---|---|---|
| *Student-level characteristics of citizenship education* | | |
| Students' perceived unsafety at school [1-4] | **1.37(0.43)** | **1.86(0.70)** |
| Students' perceived teacher-student relationship [1-5] | **3.85(0.67)** | **3.70(0.82)** |
| Students' perceived student-student relationship [1-5] | **4.27(0.62)** | **3.98(0.75)** |
| Students' perceived openness of classroom climate for discussion [1-4] | **2.41(0.54)** | **2.72(0.67)** |
| *Teacher-level characteristics of citizenship education* | | |
| Teachers' attention to personal development and social competence in citizenship education [1-5] | **4.57(0.41)** | **4.69(0.36)** |
| Teachers' attention to promoting social involvement and participation in citizenship education [1-5] | **3.69(0.47)** | **3.64(0.57)** |
| Teachers' promotion of dealing with differences and fostering of an open attitude towards differences in citizenship education [1-5] | **4.17(0.43)** | **4.21(0.45)** |
| Teachers' perceived openness of classroom climate for discussion [1-4] | **2.54(0.45)** | **2.31(0.47)** |
| The extent to which learning outcomes of citizenship education are monitored according to teachers [0-2] | **0.64(0.78)** | **0.45(0.75)** |
| The extent to which a vision on citizenship education is formulated according to teachers [1-4] | **2.19(0.98)** | **1.95(1.09)** |
| The extent to which citizenship education is translated to concrete learning goals according to teachers [1-2] | **1.29(0.32)** | **1.21(0.29)** |
| The extent to which the curriculum of citizenship education is agreed upon throughout different educational levels according to teachers [1-5] | **3.31(0.62)** | **3.07(0.53)** |
| The importance of citizenship education according to teachers [1-5] | ***4.15(0.50)*** | ***4.19(0.54)*** |

*Note*. **p Δ < 0.01 (bold)**; ***p Δ < 0.05 (bold and italic).*** Brackets indicate the answer option range.

**Table 3**

*Empty Model, Single Predictor and Control Variables with Knowledge, Applied Skills, Self-estimated Skills and Knowledge as Dependent Variables*

| | Knowledge | | | Applied skills | | | Self-estimated skills | | | Attitude | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 *B(SE)* | Model 1 *B(SE)* | Model 2 *B(SE)* | Model 0 *B(SE)* | Model 1 *B(SE)* | Model 2 *B(SE)* | Model 0 *B(SE)* | Model 1 *B(SE)* | Model 2 *B(SE)* | Model 0 *B(SE)* | Model 1 *B(SE)* | Model 2 *B(SE)* |
| **Fixed part** | | | | | | | | | | | | |
| (Intercept) | **.25(.01)** | **.35(.01)** | .11(.10) | **.29(.01)** | **.38(.01)** | .01(.13) | **.24(.01)** | **.30(.01)** | .14(.12) | **.26(.01)** | **.30(.01)** | .10(.09) |
| SNPE | | **-.19(.02)** | ***-.04(.02)*** | | **-.19(.02)** | .00(.02) | | **-.12(.01)** | **-.09(.02)** | | **-.08(.01)** | **-.06(.01)** |
| *Control variables* | | | | | | | | | | | | |
| Girl | | | ***.02(.01)*** | | | **.08(.01)** | | | **.04(.01)** | | | **.08(.01)** |
| Age | | | ***-.02(.01)*** | | | -.02(.01) | | | .01(.01) | | | .01(.01) |
| Language spoken at home | | | **.03(.01)** | | | **.04(.01)** | | | ***-.03(.01)*** | | | -.01(.01) |
| Advised educational level | | | **.08(.00)** | | | **.09(.00)** | | | **.01(.00)** | | | ***.01(.00)*** |
| IQ < 75 | | | **-.06(.02)** | | | -.02(.02) | | | -.03(.02) | | | .00(.02) |
| **Random part** | | | | | | | | | | | | |
| School-level variance | .13 | .08 | .08 | .14 | .10 | .09 | .09 | .07 | .07 | .08 | .07 | .07 |
| Residual variance | .27 | .27 | .25 | .35 | .35 | .33 | .30 | .30 | .30 | .24 | .24 | .23 |
| Total variance | .40 | .35 | .33 | .49 | .45 | .42 | .39 | .37 | .37 | .32 | .31 | .30 |
| ICC | .18 | .08 | .08 | .13 | .07 | .07 | .08 | .05 | .05 | .10 | .08 | .08 |

*Note*. **p < 0.01 (bold)**; ***p < 0.05 (bold and italic)***.

***Research Question 3: The Moderating Role of Characteristics of Citizenship Education in the Relationship between Attending SNPE and Citizenship Competences***

Tables 4 and 5 provide information on the role of student-level characteristics of citizenship education in the relationship between attending SNPE and citizenship competences. The results show that for knowledge, self-estimated skills, and attitude, the main negative effect of attending SNPE was not significantly influenced by either students' experienced unsafety in the school environment, the teacher-student relationship or the student-student relationship, or students' perceived openness of classroom climate.

For applied skills, however, the findings show that three interaction effects are significant. This means that the teacher-student and student-student relationships and students' perceived openness of classroom climate mitigated the relationship between attending SNPE and applied skills. That is, the negative main effect of attending SNPE is weaker when students experience more positive relationships with the teacher and fellow students and experience more openness in the classroom climate.

Table 6 provides information on the role of teacher-level characteristics of citizenship education in the relationship between attending SNPE and citizenship. The full models can be found in Appendices 3.3-3.6.

The results show that for knowledge, applied skills, and attitude, the negative main effect of attending SNPE is weaker if teachers pay more attention to promoting social involvement and participation. Furthermore, the negative main effect of attending SNPE on knowledge, self-estimated skills, and attitude is mitigated if teachers pay more attention to, respectively, fostering an open attitude towards differences, personal development and social competence, and indicating citizenship education to be more important.

**Table 4**

*Student-level Characteristics with Knowledge and Applied Skills as Dependent Variables*

| | Knowledge | | | | | Applied skills | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 3 *B(SE)* | Model 4.1 *B(SE)* | Model 4.2 *B(SE)* | Model 4.3 *B(SE)* | Model 4.4 *B(SE)* | Model 3 *B(SE)* | Model 4.1 *B(SE)* | Model 4.2 *B(SE)* | Model 4.3 *B(SE)* | Model 4.4 *B(SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | .00(.11) | .01(.11) | .01(.11) | .02(.11) | .01(.11) | -.05(.14) | -.02(.14) | .05(.15) | .04(.15) | .01(.14) |
| SNPE | ***-.04(.02)*** | ***-.06(.03)*** | -.06(.05) | -.06(.05) | -.06(.04) | .01(.02) | -.05(.04) | ***-.14(.06)*** | ***-.14(.07)*** | ***-.11(.05)*** |
| *Control variables* | | | | | | | | | | |
| Girl | ***.02(.01)*** | ***.02(.01)*** | ***.02(.01)*** | ***.02(.01)*** | ***.02(.01)*** | **.07(.01)** | **.07(.01)** | **.07(.01)** | **.07(.01)** | **.07(.01)** |
| Age | ***-.02(.01)*** | ***-.02(.01)*** | ***-.02(.01)*** | ***-.02(.01)*** | ***-.02(.01)*** | -.02(.01) | -.02(.01) | ***-.02(.01)*** | ***-.02(.01)*** | -.02(.01) |
| Language spoken at home | **.04(.01)** | **.04(.01)** | **.03(.01)** | **.03(.01)** | **.04(.01)** | **.04(.01)** | **.04(.01)** | ***.04(.01)*** | ***.04(.01)*** | **.04(.01)** |
| Advised educational level | **.08(.00)** | **.08(.00)** | **.08(.00)** | **.08(.00)** | **.08(.00)** | **.09(.00)** | **.09(.00)** | **.09(.00)** | **.09(.00)** | **.09(.00)** |
| IQ < 75 | **-.06(.02)** | **-.06(.02)** | **-.06(.02)** | **-.06(.02)** | **-.06(.02)** | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) |
| *Student-level characteristics of cit. ed.* | | | | | | | | | | |
| Unsafe school environment | -.01(.01) | -.02(.01) | -.01(.01) | -.01(.01) | -.01(.01) | **-.04(.01)** | **-.06(.02)** | **-.04(.01)** | **-.04(.01)** | **-.04(.01)** |
| Teacher-student relationship | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | ***.02(.01)*** | ***.02(.01)*** | -.00(.01) | .02(.01) | ***.02(.01)*** |
| Student-student relationship | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.02(.01) | -.01(.01) |
| Open classroom climate | **.03(.01)** | **.03(.01)** | **.03(.01)** | **.03(.01)** | ***.03(.01)*** | **.05(.01)** | **.05(.01)** | **.05(.01)** | **.05(.01)** | .03(.01) |
| *Interaction effects* | | | | | | | | | | |
| SNPE * Unsafe school environment | | .01(.02) | | | | | .04(.02) | | | |
| SNPE * Teacher-student relationship | | | .00(.01) | | | | | ***.04(.02)*** | | |
| SNPE * Student-student relationship | | | | .01(.01) | | | | | ***.04(.02)*** | |
| SNPE * Open classroom climate | | | | | .01(.01) | | | | | ***.05(.02)*** |
| **Random part** | | | | | | | | | | |
| School-level variance | .07 | .07 | .07 | .07 | .07 | .08 | .08 | .08 | .08 | .08 |
| Residual variance | .25 | .25 | .25 | .25 | .25 | .33 | .33 | .33 | .33 | .33 |
| Total variance | .32 | .32 | .32 | .32 | .32 | .41 | .41 | .41 | .41 | .41 |
| ICC | .08 | .08 | .08 | .08 | .08 | .06 | .06 | .06 | .06 | .06 |

*Note*. **p < 0.01 (bold)**; ***p < 0.05 (bold and italic)***.

**Table 5**

*Student-level Characteristics with Self-estimated Skills and Attitude as Dependent Variables*

| | Self-estimated skills | | | | | Attitude | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 3 *B(SE)* | Model 4.1 *B(SE)* | Model 4.2 *B(SE)* | Model 4.3 *B(SE)* | Model 4.4 *B(SE)* | Model 3 *B(SE)* | Model 4.1 *B(SE)* | Model 4.2 *B(SE)* | Model 4.3 *B(SE)* | Model 4.4 *B(SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | **-.45(.12)** | **-.45(.12)** | **-.51(.13)** | **-.45(.13)** | **-.48(.12)** | **-.51(.09)** | **-.52(.09)** | **-.48(.10)** | **-.53(.10)** | **-.52(.09)** |
| SNPE | **-.12(.02)** | **-.12(.03)** | -.03(.05) | -.11(.06 | -.06(.04) | **-.07(.01)** | ***-.07(.03)*** | **-.12(.04)** | **-.05(.05)** | -.06(.03) |
| *Control variables* | | | | | | | | | | |
| Girl | **.03(.01)** | **.03(.01)** | **.03(.01)** | **.03(.01)** | **.03(.01)** | **.07(.01)** | **.07(.01)** | **.07(.01)** | **.07(.01)** | **.07(.01)** |
| Age | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| Language spoken at home | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) |
| Advised educational level | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** | **.01(.00)** |
| IQ < 75 | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) |
| *Student-level characteristics of cit. ed.* | | | | | | | | | | |
| Unsafe school environment | **.04(.01)** | **.04(.01)** | **.05(.01)** | **.04(.01)** | **.04(.01)** | .01(.01) | .02(.01) | .01(.01) | .01(.01) | .01(.01) |
| Teacher-student relationship | **.04(.01)** | **.04(.01)** | **.04(.01)** | **.04(.01)** | **.04(.01)** | **.07(.01)** | **.07(.01)** | **.06(.01)** | **.07(.01)** | **.07(.01)** |
| Student-student relationship | **.04(.01)** | ***.04(.01)*** | **.04(.01)** | **.04(.01)** | **.04(.01)** | **.03(.01)** | **.03(.01)** | **.03(.01)** | **.04(.01)** | **.03(.01)** |
| Open classroom climate | **.10(.01)** | **.10(.01)** | **.10(.01)** | **.10(.01)** | **.11(.01)** | **.10(.01)** | **.10(.01)** | **.10(.01)** | **.10(.01)** | **.10(.01)** |
| *Interaction effects* | | | | | | | | | | |
| SNPE * Unsafe school environment | | -.00(.02) | | | | | -.00 (0.01) | | | |
| SNPE * Teacher-student relationship | | | -.02(.01) | | | | | .01(.01) | | |
| SNPE * Student-student relationship | | | | -.00(.01) | | | | | -.01(.01) | |
| SNPE * Open classroom climate | | | | | -.02(.02) | | | | | -.01(.01) |
| **Random part** | | | | | | | | | | |
| School-level variance | .06 | .06 | .06 | .06 | .06 | .05 | .05 | .05 | .05 | .05 |
| Residual variance | .29 | .29 | .29 | .29 | .29 | .21 | .21 | .21 | .21 | .21 |
| Total variance | .35 | .35 | .35 | .35 | .35 | .26 | .26 | .26 | .26 | .26 |
| ICC | .04 | .04 | .04 | .04 | .04 | .06 | .06 | .06 | .06 | .06 |

*Note*. **p < 0.01 (bold)**; ***p < 0.05 (bold and italic)***.

**Table 6**

*Overview of Interaction Effects of Teacher-level Characteristics and Attending SNPE on Dependent Variables*

| | | Knowledge | Applied skills | Self-estimated skills | Attitude |
|---|---|---|---|---|---|
| | | *B(SE)* | *B(SE)* | *B(SE)* | *B(SE)* |
| Attention to personal development and social competence | SNPE | -.23(.14) | .05(.18) | ***-.40(.15)**** | -.15(.13)* |
| | Main effect | -.02(.02) | -.00(.03) | ***-.06(.02)*** | ***-.04(.02)**** |
| | Interaction | .05(.03) | -.01(.04) | ***.07(.03)*** | .02(.03) |
| Attention to promoting social involvement and participation | SNPE | **-.23(.08)** | ***-.25(.10)*** | -.10(.09)* | ***-.21(.08)**** |
| | Main effect | -.02(.02) | -.04(.02) | .00(.02) | -.03(.02) |
| | Interaction | ***.06(.02)*** | ***.07(.03)*** | .01(.02) | ***.04(.02)*** |
| Fostering an open attitude towards differences | SNPE | ***-.26(.11)*** | -.19(.14) | -.00(.12)* | .02(.10)* |
| | Main effect | ***-.06(.02)*** | -.03(.03) | .04(.02) | ***.05(.02)**** |
| | Interaction | ***.06(.03)*** | .05(.03) | -.02(.03) | -.02(.02) |
| Open classroom climate as perceived by teachers | SNPE | -.08(.06) | -.11(.08) | -.06(.07)* | -.05(.06)* |
| | Main effect | .03(.02)* | -.00(.02) | .03(.02) | .03(.02)* |
| | Interaction | .02(.02) | .05(.03) | -.01(.03) | .00(.02) |
| Monitoring of learning outcomes of citizenship education | SNPE | -.03(.02) | .00(.02) | **-.06(.02)*** | **-.05(.02)*** |
| | Main effect | -.01(.01) | -.00(.01) | .01(.01) | -.00(.01) |
| | Interaction | .01(.01) | .02(.02) | -.02(.02) | .00(.01) |
| Formulation of a vision in citizenship education | SNPE | -.00(.03) | .04(.04) | **-.12(.03)*** | ***-.06(.03)**** |
| | Main effect | .00(.01) | .00(.01) | .01(.01)* | .01(.01) |
| | Interaction | -.01(.01) | -.01(.01) | .02(.01) | .01(.01) |
| Translation of citizenship education into learning goals | SNPE | .03(.05) | .03(.06) | -.11(.05)* | -.08(.05)* |
| | Main effect | .02(.03) | .03(.04) | -.04(.03) | -.01(.03) |
| | Interaction | -.04(.04) | -.02(.05) | .02(.04) | .02(.03) |
| Agreement upon the curriculum of citizenship education | SNPE | .03(.07) | -.00(.09) | -.14(.07)* | -.02(.06)* |
| | Main effect | .02(.01) | .02(.02) | -.02(.01) | -.01(.01) |
| | Interaction | -.01(.02) | .00(.03) | .02(.02) | -.01(.02) |
| Importance of citizenship education according to teachers | SNPE | -.12(.09) | -.01(.12) | -.13(.10)* | **-.28(.08)*** |
| | Main effect | -.00(.02) | .02(.02) | -.00(.02) | -.03(.01) |
| | Interaction | .02(.02) | .00(.03) | .01(.02) | **.06(.02)** |

*Note*. **p < 0.01 (bold)**; ***p < 0.05 (bold and italic)***. An asterisk (*) indicates that an effect was significant in the previous model with only the predictors (and no interaction effects).

In sum, the answer to the third research question consists of four parts. First, experienced unsafety in the school environment did not moderate the relationship between attending SNPE and knowledge, applied skills, self-estimated skills, or attitude. This means that the first hypothesis is falsified. Second, the teacher-student relationship moderated the relationship between attending SNPE and applied skills but not the relationships between attending SNPE and knowledge, self-estimated skills, and attitude. The same holds for the student-student relationship and students' perceived openness of classroom climate. This means that the second, third, and fourth hypotheses are confirmed only for applied skills. At last, we expected that the hypothesized interaction effects were stronger for knowledge and applied skills than self-estimated skills and attitude. This is true in that we found significant interaction effects for applied skills. Still, it does not comply with the results for knowledge – which, like applied-skills, also consists of a cognitive component but showed no significant and stronger effects than self-estimated skills and attitude

***Alternative Interpretation of the Interaction Effects with Student-level Characteristics***

Because some hypotheses including student-level characteristics of citizenship education could not be confirmed by the findings of the third research question, we explored an alternative interpretation of the interaction effects. As the modelled interaction effects can also be interpretated with SNPE as the moderator instead of the student-level characteristics of citizenship education, this was also considered. When interpreting the interaction effects like this, the insignificant interaction effects indicate that the relations between the characteristics of citizenship education (i.e., unsafe school environment, teacher-student relationship, student-student relationship, and open classroom climate) and civic knowledge, attitude, and self-estimated skills do not differ between students who attend SNPE versus students who attend RPE. The same holds for the relationship between unsafe school environment and applied skills. In most of these non-significant interaction effects, the main effects of unsafety in the school environment, the teacher-student relationship, the student-student relationship and an open classroom climate were significant. In this respect, regardless of whether students attended SNPE or RPE, it can be concluded that:

(i) an open classroom climate was positively related to citizenship knowledge ($b$ = 0.03, $SE$ = 0.01, $p < 0.01$);

(ii) experienced unsafety in the school environment was negatively related to applied skills ($b$ = -0.04, $SE$ = 0.01, $p < 0.01$);

(iii) experienced unsafety in the school environment ($b = 0.04$, $SE = 0.01$, $p < 0.01$), the teacher-student relationship ($b = 0.04$, $SE = 0.01$, $p < 0.01$), the student-student relationship ($b = 0.04$, $SE = 0.01$, $p < 0.01$) and an open classroom climate ($b = 0.10$, $SE = 0.01$, $p < 0.01$) were positively related to self-estimated skills;

(iv) the teacher-student relationship ($b = 0.07$, $SE = 0.01$, $p < 0.01$), student-student relationship ($b = 0.03$, $SE = 0.01$, $p < 0.01$) and an open classroom climate ($b = 0.10$, $SE = 0.01$, $p < 0.01$) were positively related to citizenship attitude.



In contrast, for applied skills, the interaction effects between the teacher-student relationship, student-student relationship and open classroom climate and attending SNPE were significant. A post-hoc analysis of the simple slopes indicated that:

(i) the simple slope of the teacher-student relationship was $b = -0.00$, $SE = 0.01$, $p = 0.97$ when students attended RPE, and $b = 0.04$, $SE = 0.01$, $p < 0.01$ when students attended SNPE;

(ii) the simple slope of the student-student relationship was $b = -0.02$, $SE = 0.01$, $p = 0.05$ when students attended RPE, and $b = 0.01$, $SE = 0.01$, $p = 0.28$ when students attended SNPE;

(iii) the simple slope of an open classroom climate was $b = 0.03$, $SE = 0.01$, $p = 0.07$ when students attended RPE, and $b = 0.07$, $SE = 0.01$, $p < 0.01$ when students attended SNPE.

In sum, when attending SNPE is seen as the moderator, it can be concluded that the relationships between the teacher-student relationship and applied skills, and between an open classroom climate and applied skills, are positive and significant when students attended SNPE. However, in most cases, the relationship between characteristics of citizenship education and citizenship competences of students did not change based on whether students attended SNPE.

## Discussion

Several recent studies focused on characteristics of effective citizenship education in primary and secondary education (Coopmans et al., 2020; Dijkstra et al., 2015; Wanders, Van der Veen et al., 2020). Identified aspects are the teacher-student and student-student relationship and students' perceived openness of classroom climate (Geboers et al., 2013; Wanders, Dijkstra et al., 2020). In SNPE, insight into the learning outcomes and characteristics

of citizenship education is largely missing. Hence, this study examined differences in citizenship competences between students attending SNPE versus RPE and the role of characteristics of citizenship education.

The results showed that, compared to students attending RPE, students attending SNPE experienced more unsafety in the school environment and lower quality of the teacher-student and student-student relationships. This seems to match previous research demonstrating that students in SNPE generally were more likely to engage in bullying behaviour (De Sousa et al., 2021), were more at risk of encountering difficulties in forming affective relationships with teachers (Baker et al., 2008; Freire et al., 2020; Zee & Roorda, 2018) and students (Carter et al., 2014; Pinto et al., 2019; Schwab, 2019), and experienced more difficulty with their socio-emotional competences as compared to students in RPE (Frostad & Pijl, 2007). Furthermore, our findings demonstrate that, whereas teachers in SNPE experienced the classroom climate to be less open as opposed to teachers in RPE, students in SNPE experienced more openness in the classroom climate than students in RPE. Students' experience seems to be in contrast with the results we have found for experienced unsafety in the school environment and the teacher-student and student-student relationship, which, at least based on empirical findings regarding regular primary and secondary education, can be considered important prerequisites for an open classroom climate.

Concerning other aspects of citizenship education, teachers in SNPE indicated less attention to promoting social involvement and participation, less monitoring of outcomes, less formulation of a vision, less concrete learning goals, and less agreement in the school team on the curriculum. In contrast, they indicated more attention to personal development and social competences, more promotion of dealing with differences and promoting an open attitude towards differences; and more importance attributed to citizenship education. These findings suggest that citizenship education in SNPE is not very systematic and results-oriented (cf. Hoek et al., 2022) but, rather, focuses on the challenges regarding social competence and interpersonal relationships that students in SNPE may daily encounter.

Concerning the difference in citizenship competences of students, our findings showed that students in SNPE generally obtained lower scores on knowledge, self-estimated skills, and attitude than students in RPE. For applied skills, it did not matter whether students attended SNPE or RPE. A possible explanation for lower scores on self-estimated skills and attitude may be the difficulty that students in SNPE encounter with social skills, such as self-confidence in estimating one's skills (Douma et al., 2022; Van der Veen et al., 2010) or the ability to see things from different perspectives (Conson et al., 2015; O'Kearney et al., 2017). In addition,

students in SNPE may have encountered more difficulty with knowledge questions because they require knowledge of general and distal situations. In contrast, applied skills require students to use existing knowledge to solve questions about everyday situations close to their living world. These findings align with a previous study that measured knowledge, self-estimated skills, and attitude among Dutch students in RPE and SNPE (Ledoux et al., 2012). These authors found that students with severe learning or behavioural difficulties attending SNPE also obtained lower scores on knowledge, self-estimated skills, and attitude. The design of this study differentiated between students with severe and moderate learning or behavioural difficulties, and students with moderate learning or behavioural difficulties obtained higher scores on self-estimated skills and attitude than students attending RPE. While we could not make this difference, this underscores that in future research, it is important to distinguish between the characteristics of learning and behavioural difficulties.



Contrary to our hypotheses, we found no evidence that the teacher-student and student-student relationship and students' perceived openness of the classroom climate altered the negative relationship between attending SNPE and knowledge, self-estimated skills, and attitude. While there seems to be no probable explanation for these results, it must be noted that students in SNPE do not form a clearly defined group (Pijl et al., 2008) and cope with very different educational challenges (Van der Veen et al., 2010). Moreover, the RPE and SNPE studies occurred during the COVID pandemic. School closure, social distancing and increased stress and anxiety may have affected citizenship competences of students in RPE and SNPE, with possibly stronger effects for students in SNPE as they rely more on care, support and structure that the school environment usually offers them (Hurwitz et al., 2022; Yakut, 2021). Also contrary to our hypothesis was the fact that experienced unsafety in the school environment did not change any of the relationships between attending SNPE and knowledge, applied skills, self-estimated skills, and attitude. A possible explanation may be that disruptive behaviour causing unsafety in the school environment is more prevalent among students in SNPE (Farmer et al., 2015; Ledwell & King, 2015; Rose & Espelage, 2012). As a result, unsafety in the school environment may be considered more 'normal' than in RPE and, therefore, 'numb out' the possible effect on the relationship between attending SNPE and citizenship competences. At last, in line with our hypothesis, a positive teacher-student and student-student relationship, as well as an open classroom climate as perceived by students, mitigated the negative relationship between attending SNPE and applied skills.

As some of the hypotheses could not be confirmed by the findings of the moderation analyses, alternative interpretations of the interaction effects (i.e., with SNPE treated as

moderator, instead of characteristics of citizenship education) were exploratorily examined. When SNPE is seen as the moderator, the significant interaction effects indicate that the relationship between three characteristics of citizenship education (the teacher-student relationship, the student-student relationship and an open classroom climate) and applied skills is different based on whether students attend SNPE. More specific, probing the interactions pointed out that the relationship between two characteristics (i.e., the teacher-student relationship and an open classroom climate) and applied skills is positive and significant when students attended SNPE. However, in most cases, the relationship between characteristics of citizenship education and citizenship competences (predominantly citizenship knowledge, attitude, and self-estimated skills) of students did not change based on whether students attended SNPE.

We also performed exploratory analyses on the role of additional characteristics of effective citizenship education in the relationship between attending SNPE and citizenship competences. The results show that attention to promoting social involvement and participation weakened the negative relationship between attending SNPE and knowledge, applied skills, and attitude. Attention to two other specific learning goals (i.e., attention to personal development and social competence and fostering an open attitude towards differences) improved the relationship between attending SNPE and, respectively, self-estimated skills and knowledge. The fact that attention to these specific learning goals mitigated the negative effects of attending SNPE on various citizenship competences suggests that the attention and effort of the teacher are important conditions in the acquisition of citizenship competences in SNPE.

To conclude, this study contributes to insight into the role of characteristics of effective citizenship education in SNPE versus RPE. Whereas the role of some characteristics was not as expected and ask for further research, this study shed light on several important factors, such as the attention for specific learning goals in citizenship education, that form an important basis for further research into characteristics of effective citizenship education in SNPE.

# Chapter 4

# Young Students' Attitudes Towards Fundamental Democratic Values: The Construction of a Measurement Instrument


## Abstract

Schools carry the important task of fostering fundamental democratic values (FDVs) like equality and tolerance in students. Insight into students' learning outcomes concerning FDVs helps to improve the quality of citizenship education. Hence, this study examined to what extent a valid and reliable measurement instrument can be developed to gain insight into young students' attitude towards FDVs, and what can be concluded based on this measurement.

We constructed the Fundamental Democratic Values Questionnaire (FDVQ) based on existing conceptualisations of FDVs. After that, we conducted two pilot studies (N = 302 and N = 227) and a first measurement (N = 3278) among Dutch students in grades 5 and 6.

The results support the validity and reliability of the FDVQ to measure young students' attitude towards FDVs. Moreover, the findings indicated that support for FDVs among students was generally high. These findings and the implications of the FDVQ are further discussed.



*Keywords*: Fundamental democratic values; Citizenship education; Civic attitude; Instrument construction.





*This chapter is based on: Hoek, L., Munniksma, A., & Dijkstra, A. B. (2023). Young students' attitudes towards fundamental democratic values: The construction of a measurement instrument. *Manuscript revised and resubmitted.*

## Introduction

In pluriform societies, democracy is a form of government that supports the humane and structured co-living of (groups of) individuals who are different in terms of religion, culture, political ideas, and so on (De Winter, 2004). It does so by being responsive to the majority's views while providing space to and protecting the rights of minorities (Council of Europe, 2018). For democracy to persevere, it is important that citizens have an interest and active involvement in the democratic society and underscore and have knowledge of fundamental democratic values (FDVs) like equality, tolerance and non-discrimination (Levinson, 1999). In addition, citizens of pluriform democracy must also be able to make critical contributions (Wardekker, 2001), for example, by critically exploring strategies, perspectives and situations of (in)equality and (in)justice (Westheimer, 2008). Without actively involved and informed citizens who comply with such a 'democratic way of life' (cf. Dewey, 1923), democracy becomes fragile. A democratic way of life presumes citizens to acknowledge and accept that others may have different interests and to be willing to deal with conflicts via dialogue and negotiation or, ultimately, the rule of law. Democratic societies benefit in terms of resilience when (young) citizens live up to a democratic way of life (Dekker & Den Ridder, 2016; Welzel & Dalton, 2014).

The importance of FDVs, such as equality, non-discrimination and respecting the rule of law, for the resilience of a pluriform democracy is underscored by the fact that they are anchored in many national constitutions and the Universal Declaration of Human Rights (UDHR) (*Universal Declaration of Human Rights*, 1948). Among citizens, the importance of underscoring FDVs is also acknowledged. Using decades of World Survey Data, Welzel (2021), for example, argued that support for so-called emancipative values (i.e., a subset of values that combines an emphasis on freedom of choice and equality of opportunities) among citizens throughout the world is on the rise, leading to a stronger commitment to democratic principles among younger generations of citizens. In addition, research shows that adult citizens perceived various democratic values (e.g., knowledge of values and norms, rules of conduct, and human rights, including children's rights) to be important for young citizens to learn (Dekker & Den Ridder, 2016; Van Goethem et al., 2020). Despite their importance, the acceptance and underscoring of FDVs by individual citizens is not a given. Or, as White (1999) puts it: 'democrats are made, not born' (p. 59).

Instead, FDVs require to be educated, promoted and maintained. Since teaching students the competences needed to partake in a democratic society is an important goal of education (Dewey, 1923), and education is a common socialisation context shared by all young

citizens (Beane, 2013), it is important that schools carry out the task of fostering these competences in students (Council of Europe, 2018). In line with that, as a response to societal changes such as increased diversity, individualisation, and polarisation (Mattei & Broeks, 2018; Organisation for Economic Co-operation and Development, 2010; Tyler & Iyengar, 2022), many countries throughout Europe have appealed to schools with the task of fostering democratic citizenship competences in students via citizenship education (Eurydice, 2017). Citizenship competences are generally understood as the knowledge, attitude and skills young citizens need to participate in a democratic and pluriform society (Schulz et al., 2018). To gain citizenship knowledge and skills, the underscoring of FDVs can be seen as an important basis: as a student, you have to be open, motivated and willing to learn about the perspectives of others and let others learn from your perspectives (De Winter, 2004; White, 1999).



In summary, FDVs are widely conceived as important for democratic and pluriform societies, require active fostering (i.e., 'democrats are made, not born'), and (citizenship) education is one of the socialising agents that plays a role in this. Even though there is no fixed or demarcated set of FDVs that should be promoted in education (e.g., differences may reveal between a liberal and communitarianism perspective, also see Eidhof et al., 2016; Miller, 2000), common ground can be found in the *regula aurea* or Golden Rule (Wattles, 1996). The Golden Rule is a widely held philosophical and religious principle better known as 'to treat others as you want others to treat you'. As such, this norm of reciprocity can be seen as an important foundation for value formation and may help schools to reflect on the commonly held values they want to teach.

To gain insight into the quality and effectiveness of education focused on promoting FDVs among students, it is important to have insight into students' learning outcomes. A standardised measurement instrument that provides insight into students' attitudes towards FDVs helps schools adjust their education to what students need. Although instruments for scientific use are available, no measurement instrument for schools is currently available with the main goal of capturing young students' attitudes towards FDVs. Hence, the present study aims to examine to what extent a reliable and valid measurement instrument suitable for educational practice can be developed to gain insight into students' attitudes towards FDVs. The following research question is central in this study: 'How can young students' attitude towards fundamental democratic values be reliably and validly measured, and what can be concluded based on this measurement?'

## Theoretical Framework

The conceptualisation of FDVs received consistent attention in the scholarly literature. According to Beane (2013), the competences needed for participation in a democratic society include that students become aware of values such as respect for human dignity, equity, freedom, and social responsibility. Moreover, it includes that students become competent in skills such as critical thinking, problem-solving, collaboration, information and data gathering, reflection, and participatory planning. Eidhof et al. (2016) raise the label consensus goals, which are described as general democratic goals that reflect what the majority of a population perceives as important democratic principles to foster in youth. The authors have synthesised the literature and argue that consensus goals generally consist of three main democratic values: tolerance of diversity (Barber, 1984; Galston, 2001; Kymlicka, 1989; Van Gunsteren, 1998), non-violence (Galston, 2001), and civic engagement (Almond & Verba, 1963; Putnam, 2000). In addition, Almond and Verba (1963) argued that, for democratic societies to last, it is not only important for FDVs and attitudes in the political and public sphere (e.g., support for democratic institutions, acceptance of policy decisions) to be anchored in society, but also for FDVs that belong to the social and cultural roots of democracy, such as tolerance, rejection of violence, sense of belonging and solidarity towards a community, openness towards individuals with other faiths and beliefs. In the UK context, Janmaat (2018) examined the extent to which young people supported so-called 'fundamental British values', which the UK Department of Education (2014) defined as democracy, the rule of law, individual liberty, and mutual respect and tolerance of those with different faiths and beliefs. Even though labelling these beliefs as 'fundamentally British' is criticised for being a misspecification (cf. Lander, 2016; Tomlinson, 2015), the values as such can be considered fundamental for democratic societies and show resemblance with the values mentioned in national constitutions and in documents such as the UDHR.

In line with their importance for the non-violent and cooperative co-living of different (groups of) individuals in a pluriform democracy, various FDVs are captured in the UDHR. In the UDHR, among others, non-discrimination, non-violence, freedom of speech, thought and religion, equality, personal integrity, and political participation are included (*Universal Declaration of Human Rights*, 1948). Moreover, the UDHR is alleged to be a universal document with universal compliance regardless of individuals' sex, culture, ethnicity, religion, political ideas, or other. The European Convention on Human Rights (ECHR) (*European Convention on Human Rights*, 1953) and United Nations Convention on the Rights of the Child (UNCRC) (*United Nations Convention on the Rights of the Child*, 1989) both show great

similarities concerning the incorporated values with the UDHR but are, respectively, oriented towards member states of the Council of Europe and citizens younger than 18 years old.

Besides their role in the ECHR, the Council of Europe (Council of Europe, 2018) identified, among other competences, three main values and six attitudes in the Reference Framework of Competences for Democratic Culture (RFCDC). In this respect, values are 'general beliefs that individuals hold about the desired goals that should be striven for in life' (Council of Europe, 2018, p. 38). Their normative, prescriptive quality characterises them. Attitudes are 'overall mental orientations which an individual adopts towards someone or something' (Council of Europe, 2018, p. 41). The RFCDC was designed to help member states promote a democratic culture among its citizens by providing a clear focus and understanding of common goals in citizenship education. The values in the RFCDC are human dignity and human rights; cultural diversity; and democracy, justice, fairness, equality and the rule of law (Council of Europe, 2018). The attitudes in the RFCDC are openness to cultural otherness and other beliefs, world views and practices; respect; civic-mindedness (i.e., an attitude of belonging to, mindfulness of, a sense of solidarity with, interest in, willingness to contribute to, commitment to, and a sense of accountability towards a larger community or social group); moral responsibility; self-efficacy; and tolerance of ambiguity (i.e., an attitude towards objects, events and situations which are perceived to be uncertain and subject to multiple conflicting or incompatible interpretations) (Council of Europe, 2018).

In sum, it can be concluded that both the scientific reasoning as well as highly influential policy documents mentioned (aspects of) being open, tolerant and respectful towards individuals with other beliefs, world views, faiths and practices and, related, the principles of non-discrimination, non-violence and equality as fundamental to democracy. In addition, in both lines, (aspects of) human dignity, personal integrity, individual liberty, and freedom (of speech, thought and religion) are mentioned. The same holds for human rights, justice, fairness, the rule of law, support for democratic institutions, and acceptance of policy decisions. At last, (aspects of) being part of a larger community or social group were mentioned, such as solidarity, accountability and (moral and social) responsibility towards this group.

**The Present Study**

In many countries, education is appointed by law to foster FDVs in young citizens. This is, for example, visible in the UK (Janmaat, 2018), in the Netherlands, Norway, Scotland, and Sweden (Dijkstra & De la Motte, 2014) and many other countries throughout Europe (Eurydice, 2017) and around the world (Schulz et al., 2018). The present study focused on the

Netherlands, where the national government identified eight FDVs that should be fostered in citizenship education, i.e.: equality, understanding of others, freedom of speech, autonomy, tolerance, the rejection of intolerance, the rejection of discrimination, and sense of responsibility (Ministry of Education, Culture and Science, 2021). The Dutch government describes these eight FDVs as universal values that are widely legitimated as visible in the national constitution and the UDHR. Moreover, the statutory citizenship obligation for schools, including the task of promoting these FDVs, leaves autonomy with schools to determine what values they want to foster or emphasise in their students but, at the same time, underscore that these eight values are essential for participation in a pluriform and democratic society – like the Dutch society – and, hence, should not be undermined in education. To adequately prepare young citizens for participation in society, the minimum requirement to ask from schools is to actively foster these FDVs in students and ensure that the education they give does not conflict with them.

While the set of eight FDVs as specified by the Dutch government is not exhaustive, it resembles universal values given its great overlap with conceptualisations of FDVs in the scholarly literature (Almond & Verba, 1963; Beane, 2013; Eidhof et al., 2016; Janmaat, 2018) and international policy documents such as the UDHR, ECHR, UNCRC, and RFCDC. In determining what values should be included in the Dutch legislation, the legislator took their relevance for and connection to the development phase of young students. In that respect, the FDV of equality was, for example, considered more overarching than the value of ensuring that all citizens have a fair judicial trial. Moreover, the FDV of equality was considered to be a better fit for the experience and lives of young students than ensuring a fair trial for all citizens. Although values such as the latter were not directly included in the legislation, they are still indirectly addressed. Likewise, non-specified values like freedom of thought, consciousness and religion are, for example, indirectly addressed in the value of freedom of speech. The remainder of this paper describes the construction of a measurement instrument that can be used to gain insight into students' attitudes towards equality, understanding of others, freedom of speech, autonomy, tolerance, the rejection of intolerance, the rejection of discrimination, and a sense of responsibility. As such, this measurement instrument was designed to align with the context in which the study took place and followed the conceptualisation of FDVs as visible in the legislation that applies to Dutch schools.

## Method

This section consecutively describes the participants, materials and procedure of two pilot studies and one first measurement conducted to construct the Fundamental Democratic Values Questionnaire (FDVQ).

### Participants

The scope of this paper was delineated to students in grades 5 and 6 of regular primary education (RPE) and students in the final ('leaving') grade of special needs primary education (SNPE)[6]. In the first pilot study, we used data from 122 students from grade 5 in RPE (53% boys), 150 from grade 6 in RPE (47% boys) and 30 students from SNPE (83% boys). In the second pilot study, we used data from 125 students from grade 5 in RPE (46% boys) and 102 from grade 6 in RPE (49% boys). In the first measurement using the final version of the FDVQ, we used data from 1367 students from grade 5 in RPE (50% boys), 1606 from grade 6 in RPE (47% boys) and 305 students from SNPE (69% boys). In this first measurement, 7% of the students had a migration background, and 19% indicated that their home language was other than Dutch, Frisian or a regional Dutch language. The descriptive statistics of the samples in the two pilot studies and the first measurement using the final instrument can be found in Table 1.

**Table 1**

*Characteristics of Participants*

| | Educational sector | Grade | N | Sex | Mean age |
|---|---|---|---|---|---|
| Pilot 1 | Regular primary education | 5 | 122 | 53% boys | 9.79 |
| | | 6 | 150 | 47% boys | 10.69 |
| | Special needs primary education | 5/6 | 30 | 83% boys | 10.87 |
| Pilot 2 | Regular primary education | 5 | 125 | 46% boys | 10.4 |
| | | 6 | 102 | 49.0% boys | 11.5 |
| Measurement 1 | Regular primary education | 5 | 1367 | 50% boys<br>50% girls<br><0.1% other/I'd rather not say | 10.77 |
| | | 6 | 1606 | 47% boys<br>50.7% girls | 11.72 |

[6] The Fundamental Democratic Values Questionnaire was also administered to students in secondary education and demonstrated comparable results. However, an elaborate discussion of the findings did not fit within the scope of this paper.

| | | | | | |
|---|---|---|---|---|---|
| | | | | <0.1% other/I'd rather not say | |
| | Special needs primary education | 5/6 | 305 | 68.5% boys<br>29.2% girls<br><0.1% other/I'd rather not say | 12.04 |

**Materials**

The FDVs, as conceptualised in the Dutch legislation and in this study, consist of equality (i.e., all individuals are of equal value, despite beliefs or faith), understanding of others (i.e., understanding why individuals or communities have certain beliefs or faiths), freedom of speech (i.e., all individuals are allowed to say and write what they think or go against conceptions of others, as long as it is in accordance with the law), autonomy (i.e., all individuals may decide for themselves who they want to be and how they want to live their live, as long as it is in accordance with the law), tolerance (i.e., accepting and providing space to the beliefs, faiths or behaviour of others even if you do not agree with it, as long as it is in accordance with the law), rejection of intolerance (i.e., rejecting behaviour in which other individuals beliefs, faiths or behaviours are not accepted or allowed to take space), rejection of discrimination (i.e., rejecting behaviour in which other individuals are being discriminated against based on their beliefs, faiths or behaviours), and sense of responsibility (i.e., taking responsibility for what you say or do, and for what you do not say and do not do, and taking into account what that means for others, while trying not to harm others and help democracy to function well) (Dijkstra, El Khayati, et al., 2014).

To establish the FDVQ to measure students' attitude towards FDVs, the FDVs originating from the Dutch law on citizenship education (cf. Ministry of Education, Culture and Science, 2021) were carefully operationalised into eight Likert-type scales. Each item had four answer options: 'totally disagree', 'somewhat disagree', 'somewhat agree', and 'totally agree'. This was decided to force students to choose between a more negative ('totally disagree'; 'somewhat disagree') or positive ('somewhat agree'; 'totally agree') answer. Even though research is inconclusive on whether or not to include a neutral midpoint (e.g., Chyung et al., 2017), scholars have posed that excluding a neutral midpoint minimises socially desirable responses of respondents (Garland, 1991; Johns, 2005), which is important to take into account when studying citizenship competences of students (Ten Dam et al., 2013).

In addition to their attitude towards eight FDVs, information on students' sex, age, grade, educational sector, home language, and migration background was gathered. For home

language, students indicated whether their home language was Dutch; Frisian or a Dutch regional language; Turkish; English; Moroccan or Berber; Surinamese or an Antillean language; or something different. For migration background, students indicated whether they lived in the Netherlands for less than a year; one to three years; four to five years; longer than five years; or whether they have always lived in the Netherlands.

**Procedure**

As part of the construction phase, two pilot studies were conducted in different groups of students: one between April and June 2021; one between February and March 2022. After each pilot, the psychometrics of the proposed items and scales were checked by examining descriptive statistics (e.g., mean, standard deviation, answer pattern, skewness, kurtosis), Cronbach's reliability, and exploratory and confirmatory factor analyses. Furthermore, students, teachers, school leaders, researchers, and content specialists were consulted before and after each pilot. Based on the psychometrics and consultation, the items were improved into a final version of the FDVQ (see Table 2). The changes consisted mostly of maximising linguistic accessibility by simplifying the language and shortening items as much as possible. Additionally, in the final questionnaire, when asked for their sex, students could also opt for 'other/I'd rather not say', whereas, in the pilot studies, this option was not available. The final version of the FDVQ was used for the first time between March and July 2022 as part of an annual measurement of citizenship knowledge, attitude, and skills using the Citizenship Competences Questionnaire (Ten Dam et al., 2011), in which schools could voluntarily participate.

**Table 2**

*Example Items in the Final Measurement Instrument*

| Fundamental Democratic Values | Example item |
|---|---|
| Equality | All people are equal, including people who live very differently from me. |
| Understanding of others | When somebody has a different opinion than I do, I try to understand why. |
| Freedom of speech | If you disagree with the government, you are allowed to say so. |
| Autonomy | You may decide for yourself whether you are in love with a boy or a girl. |
| Tolerance | It is important to respect people with a different opinion than mine. |
| Rejection of intolerance | I think it is wrong when somebody is not respected because of their religion. |
| Rejection of discrimination | I think it is wrong when somebody is discriminated against because of their disability. |
| Sense of responsibility | People should try not to harm others. |

Before each (pilot) measurement occurred, parents received an information letter and were allowed to deny their children's involvement. Before students filled in the questionnaire, they were ensured anonymity and allowed to quit participation at all times. After giving the instructions, students took approximately half an hour to complete the questionnaire.

**Missing Observations**

The data that were gathered as part of the first measurement after the pilot studies (using the final measurement instrument) consisted of 3373 students (< 1% missing values) from 69 schools for primary education. As the percentage of missing values was very small and manifested more towards the end of the questionnaire, incomplete cases were deleted from the dataset instead of using imputation methods. This resulted in a dataset consisting of 3278 students.

## Results

This section reports the findings of several psychometric analyses based on advised thresholds. As such, it provides information on the quality of the instrument, such as an analysis of descriptive statistics (e.g., mean, standard deviation, answer pattern, skewness, kurtosis), correlation analyses, reliability analyses, the factor structure of the exploratory factor analyses, fit indices of the confirmatory factor analyses, an assessment of measurement invariance, response frequencies, and relationships with external constructs.

**Descriptive Analyses**

Pearson correlation coefficients between the eight scales were examined to provide insight into the relations between FDVs (Table 3). What stands out is that the correlations are all positive and predominantly high. Furthermore, the correlation coefficients between 'rejection of intolerance' and 'rejection of discrimination' and the other FDVs, particularly 'understanding of others' and 'freedom of speech', are somewhat lower than the other correlations.

**Table 3**
*Correlation Coefficients*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. |
|---|---|---|---|---|---|---|---|---|
| 1. Equality | 1.00 | | | | | | | |
| 2. Understanding of others | 0.44 | 1.00 | | | | | | |
| 3. Freedom of speech | 0.40 | 0.31 | 1.00 | | | | | |
| 4. Autonomy | 0.62 | 0.32 | 0.47 | 1.00 | | | | |
| 5. Tolerance | 0.61 | 0.40 | 0.47 | 0.68 | 1.00 | | | |
| 6. Rejection of intolerance | 0.39 | 0.26 | 0.24 | 0.30 | 0.35 | 1.00 | | |
| 7. Rejection of discrimination | 0.43 | 0.21 | 0.26 | 0.35 | 0.40 | 0.54 | 1.00 | |
| 8. Sense of responsibility | 0.54 | 0.43 | 0.37 | 0.50 | 0.62 | 0.37 | 0.45 | 1.00 |

*Note.* Correlations between .1 and .3 are considered low; correlations between .3 and .5 are considered medium; correlations between .5 and 1.0 are considered high.



Table 4 depicts the reliability coefficients, the number of items, the mean scale scores, standard deviations, and values for skewness and kurtosis. For all scales, values for Cronbach's alpha were above the advised threshold of 0.70 (Cortina, 1993). The mean scores point to the manifestation of a ceiling effect, which is unsurprising since the students who participated in this study live in a society where FDVs are generally broadly shared (Van Goethem et al., 2020). Furthermore, values between -2 and +2 for skewness and -7 and +7 for kurtosis indicate a normal distribution (Byrne, 2010; Hair et al., 2010). Some values for skewness are smaller than -2, which underscores the importance of choosing a robust maximum likelihood as an estimator in the confirmatory factor analyses to account for this nonnormality (Knekta et al., 2019).

**Table 4**

*Reliability Coefficients (Cronbach's Alpha), Number of Items (N), Mean Scale Scores, Standard Deviations (sd), and Skewness and Kurtosis*

| | Alpha | N | Mean | Sd | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| Equality | 0.83 | 4 | 3.64 | 0.57 | -1.85 | 3.43 |
| Understanding of others | 0.85 | 4 | 3.15 | 0.68 | -0.69 | 0.28 |
| Freedom of speech | 0.71 | 5 | 3.46 | 0.52 | -1.23 | 2.14 |
| Autonomy | 0.77 | 4 | 3.67 | 0.53 | -2.01 | 4.20 |
| Tolerance | 0.80 | 4 | 3.67 | 0.49 | -1.94 | 4.35 |
| Rejection of intolerance | 0.92 | 4 | 3.43 | 0.83 | -1.67 | 1.94 |
| Rejection of discrimination | 0.95 | 4 | 3.70 | 0.70 | -2.67 | 6.59 |
| Sense of responsibility | 0.83 | 5 | 3.61 | 0.52 | -1.88 | 4.32 |

**Exploratory Factor Analysis**

Exploratory factor analysis (EFA) was performed to examine how well the data fit the intended eight-factor structure. This was done by appointing eight factors and examining on

what factor each item has the highest factor loading using varimax rotation. The results (Appendix 4.1) reveal that the data follow the intended factor structure of eight distinguishable FDVs. In two cases, the intended scale does not entirely group together: for 'equality', the item 'Everybody is equal, including men who like men, women who like women, and men who like women' has a higher factor loading (0.499) outside the intended scale. In addition, for 'autonomy', the item 'You may decide for yourself whether you are in love with a boy or a girl' also has a higher factor loading (0. 884) outside the intended scale. Both items are about sexual diversity and group better together. In addition, the scales 'autonomy' and 'tolerance' have the highest factor loadings on the same factor. This may be the cause because the items in 'autonomy' are about having the right to make your own decisions, whereas in 'tolerance', the items are about respecting others to make their own decisions.

**Confirmatory Factor Analysis**

Confirmatory factor analyses (CFA) examine how well the data fit the intended factor structure. As opposed to the procedure in EFA, the items are fixed in the intended scales in CFA. As Kline (2010) recommended, it was assessed whether the model fit was adequate before interpreting the model estimates. This was done by looking at the comparative fit index (CFI) and root mean square error of approximation (RMSEA). For CFI, a value above 0.95 is preferred (Hu & Bentler, 1999); for RMSEA, a value lower than 0.050 indicates close fit (Browne & Cudeck, 1993).

First, a first-order factor model with 34 items as indicators of 8 uncorrelated FDVs was fitted to the data and yielded an unsatisfactory model fit ($\chi^2(527) = 8067.417$, $N = 3278$, $p < .001$, $CFI = 0.782$, $RMSEA = 0.089$, 90% CI [0.088, 0.091]). Therefore, the same model was fitted in which the 8 FDVs were allowed to correlate. This yielded satisfactory to close model fit ($\chi^2(499) = 2310.117$, $N = 3278$, $p < .001$, $CFI = 0.948$, $RMSEA = 0.045$, 90% CI [0.043, 0.047]) and fitted the data significantly better than the first-order factor model with uncorrelated FDVs ($p \leq 0.001$).

Despite the satisfactory to close fit of the previous model, it was explored whether a second-order factor model with 34 items as indicators of 8 FDVs as indicators of 1 latent construct 'democratic attitude' fitted the data better. This model yielded unsatisfactory to acceptable model fit ($\chi^2(526) = 2911.581$, $N = 3278$, $p < .001$, $CFI = 0.931$, $RMSEA = 0.050$, 90% CI [0.048, 0.052]) and fitted the data significantly worse than the first-order factor model with correlated FDVs ($p \leq 0.001$). Next, a first-order factor model was fitted with eight mean scores of FDVs as indicators of 1 latent construct 'democratic attitude'. This yielded poor model fit ($\chi^2(20) = 628.137$, $N = 3278$, $p < .001$, $CFI = 0.902$, $RMSEA = 0.123$, 90% CI [0.115,

0.131]) and also fitted the data significantly worse than the first-order factor model with correlated FDVs ($p \leq 0.001$). These results indicate that neither the second-order factor model with items as indicators of FDVs as indicators of a latent construct nor the first-order factor model with mean scores as indicators of a latent construct fitted the data better than the first-order factor model with items as indicators of correlated FDVs.

At last, as the correlation matrix (Table 3) pointed out that 'rejection of intolerance' and 'rejection of discrimination' were less correlated to the other FDVs, a second-order factor model was constructed with 34 items as indicators of 8 FDVs, in which 'rejection of intolerance' and 'rejection of discrimination' served as indicators of a separate latent construct as opposed to the remaining FDVs. This yielded satisfactory to close model fit ($\chi^2(524) = 2607.950$, $N = 3278$, $p < .001$, $CFI = 0.940$, $RMSEA = 0.047$, 90% CI [0.045, 0.049]) but fitted the data significantly worse than the first-order factor model with correlated FDVs ($p \leq 0.001$). Therefore, the first-order factor model with 34 items as indicators of 8 correlated FDVs was accepted and proceeded with in further analyses.



**Assessment of Measurement Invariance**

Multiple-group confirmatory factor analysis (MGCFA) was used to assess the measurement invariance (Meredith, 1993; Putnick & Bornstein, 2016) of the FDVQ to ensure meaningful comparisons across different groups of a population (Chen et al., 2005; Isac et al., 2019; Steinmetz et al., 2009). To do so, increasingly restricted models (i.e., configural, threshold, metric and scalar models) were fitted to the data. Achieving (partial) metric invariance is a minimal prerequisite for meaningful cross-group comparisons (Little, 2013). It justifies the comparison of latent variables and their associations across groups (Isac et al., 2019). Reaching (partial) scalar invariance justifies comparing latent means across groups (Reeskens & Hooghe, 2010). The Chi-square test of the difference in model fit ($\Delta \chi^2$) was used to test the null hypothesis of equal model fit for two models. If the *p*-value of the Chi-square test of difference is significant, it indicates a meaningful difference between the two models in comparison, and the more restricted model is rejected. In addition, partial metric and partial scalar invariance were tested for. In these partial models, some (but not all) of the equality constraints were released again. The standardised mean residuals and significance of modification indices (using a Bonferroni-adjusted alpha level) were consulted to find out what constraints needed to be released.

Research on citizenship competences examined and found robust differences based on sex, socioeconomic position, and migration background (e.g. Geijsel et al., 2012; Ireland et al., 2006; Kerr et al., 2007). Scholars and schools will likely be interested in examining whether

such cross-group differences also occur in students' attitudes towards FDVs. Hence, the measurement invariance was tested across sex, home language, educational sector, grade, and migration background. The students who answered the question about their sex with 'other/I'd rather not say' had to be excluded from the analyses because they were too small as a separate category. For sex, home language, educational sector and grade, (partial) metric invariance was reached (Appendix 4.2). This allows for the comparison of associations between the FDVs and the underlying items across boys versus girls, across students with Dutch or a dialect of Dutch as a home language versus another home language, across students in RPE versus SNPE, and across students in grade 5 versus grade 6. For migration background, scalar invariance was reached. This allows for comparing mean scores of the FDVs across students with and without a migration background (Appendix 4.3).

**Response Frequency**

The non-missing response frequency for each questionnaire item demonstrated, among others, a so-called 'ceiling effect': most students opted for 'somewhat agree' or 'totally agree' to indicate their attitude towards FDVs. Furthermore, the response frequencies of one of the items of 'freedom of speech' (i.e., 'Journalists are allowed to decide for themselves what they write about') and all four items of 'understanding of others' (e.g.., 'When someone has a very different opinion than I do, I try to understand why') somewhat stand out because the positive answer options are more equally divided between 'somewhat agree' and 'totally agree', whereas for most other questionnaire items, the share of 'totally agree' within the positive answer options is larger than 'somewhat agree'. The response frequencies for 'totally disagree' are largest for the aforementioned item of 'freedom of speech' and all items of 'rejection of intolerance' (e.g., 'I think it is wrong when somebody is not respected because of their religion').

**Relation between Fundamental Democratic Values and Civic Knowledge, Attitude, and Skill**

As a last analysis, the correlation coefficients between students' attitudes towards fundamental democratic values and civic knowledge, civic attitude, and self-estimated civic skills[7] were examined (see Table 5). Civic knowledge, attitude, and self-estimated skills are

[7] Citizenship knowledge is commonly understood as the knowing and understanding of democratic principles, social rules, conflict-solving, and cultural differences, for example: having knowledge about how democratic decisions are made. Questions were multiple-choice with three options. Students had to choose the answer they thought was best. An example is: 'All children have a right to… (a) an allowance; (b) choose who they want to live with; or (c) education. The correct answer here is c. All knowledge items were dichotomized into incorrect (0) or correct (1). Citizenship attitude is about the willingness to pursue democratic principles in one's own actions, for example: willing to take other perspectives into account, even if it means that one's self-interest

assumed to be related to students' underscoring of FDV but are measuring aspects of civic competences as operationalised and measured using the Citizenship Competences Questionnaire (CCQ) (Ten Dam et al., 2011). Students' civic attitude, as measured by the CCQ, focuses on students' attitudes towards four so-called social tasks: acting democratically, acting in a socially responsible manner, dealing with conflicts, and dealing with differences.

What stands out in Table 5 are the relatively lower correlation coefficients between 'rejection of intolerance' and 'rejection of discrimination' with civic attitude and self-estimated civic skills. Furthermore, the relatively lower correlation coefficients for 'understanding of others' and 'freedom of speech' with civic knowledge' indicate that civic knowledge does not necessarily relate to a positive attitude towards 'understanding of others' and 'freedom of speech'. For 'understanding of others', the correlation coefficients with civic attitude and self-estimated civic skills are relatively high, indicating a strong relationship between these constructs. The interpretation of these interrelations is discussed further in the discussion.



**Table 5**

*Correlation Coefficients*

| | Civic knowledge | Civic attitude | Self-estimated civic skills |
|---|---|---|---|
| 1. Equality | 0.49 | 0.47 | 0.39 |
| 2. Understanding of others | 0.24 | 0.63 | 0.55 |
| 3. Freedom of speech | 0.29 | 0.34 | 0.36 |
| 4. Autonomy | 0.46 | 0.37 | 0.37 |
| 5. Tolerance | 0.44 | 0.39 | 0.42 |
| 6. Rejection of intolerance | 0.34 | 0.28 | 0.25 |
| 7. Rejection of discrimination | 0.40 | 0.27 | 0.22 |
| 8. Sense of responsibility | 0.42 | 0.48 | 0.48 |

*Note.* Correlations between .1 and .3 are considered low; correlations between .3 and .5 are considered medium; correlations between .5 and 1.0 are considered high.

## Discussion

For democracy to persevere, its citizens must actively maintain it. Citizens can do so by underscoring fundamental democratic values (FDVs) like equality, tolerance and non-discrimination (Levinson, 1999). FDVs are widely conceived as important for democratic and

cannot be fully fulfilled. The phrasing of the attitude items is: 'How well does this statement apply to you?' An example is: 'I like knowing something about different religions.' The response options are (1) does not apply at all to me; (2) does not apply much to me; (3) applies a fair amount to me; or (4) applies completely to me. Citizenship skills are about the competence to put democratic principles to action, for example: the competence to pose one's own opinion and listen to opinions of others. The phrasing of the skills items is: 'How good are you at…?' An example is: '… finding a solution that everyone is satisfied with for a conflict?' The response options are: (1) not good at all; (2) not very good; (3) pretty good; or (4) very good.

pluriform societies and are anchored in international policy documents such as the Universal Declaration of Human Rights. However, FDVs require active fostering (cf. White, 1999), and schools are one of the socialising agents that play a role in this. To improve the quality and effectiveness of education aimed at fostering FDVs, insight into students' learning outcomes is important. Hence, this study aimed to examine to what extent a valid and reliable measurement instrument can be developed to gain insight into young students' attitude towards FDVs. This was done by examining the following research question: 'How can young students' attitude towards fundamental democratic values be reliably and validly measured, and what can be concluded based on this measurement?'

The current study set out with a theoretical synthesis of different conceptualisations of FDVs. Most conceptualisations included (aspects of) openness towards diversity in a broad sense; human dignity and freedom; support for democratic institutions and principles, policy decisions and human rights; and communality. In the Netherlands, where the present study took place, FDVs are conceptualised as equality, understanding of others, freedom of speech, autonomy, tolerance, the rejection of intolerance, rejection of discrimination, and sense of responsibility (Ministry of Education, Culture and Science, 2021). This set of FDVs shows similarities with conceptualisations of FDVs in the scholarly literature (Almond & Verba, 1963; Beane, 2013; Eidhof et al., 2016; Janmaat, 2018) and international policy documents such as the Universal Declaration of Human Rights (UDHR), the European Convention on Human Rights (ECHR), the United Nations Convention on the Rights of the Child (UNCRC), and the Reference Framework of Competences for Democratic Culture (RFCDC) from the Council of Europe (Council of Europe, 2018). Hence, the eight FDVs were used as the conceptual framework for constructing the Fundamental Democratic Values Questionnaire (FDVQ).

The results show that for each FDV, a reliable scale could be distinguished. In addition, a first-order factor model with 34 questionnaire items as indicators of 8 correlated FDVs fitted the data well. Altogether, the results provide empirical support for the validity and reliability of the measurement of FDVs in young students. In terms of interrelations between FDVs, the findings point out that most FDVs had moderate to strong correlations with other FDVs. This confirms that FDVs all tap upon a different but related part of FDVs as an overarching latent construct. Overall, 'rejection of intolerance' and 'rejection of discrimination' were less related to other FDVs. A possible explanation may be that rejecting an undemocratic value of others is slightly different because it entails both yourself and (an action towards) other individuals, as opposed to the other FDVs, which only involve an individual's compliance with democratic

values. Moreover, whereas the underlying questions of 'rejection of intolerance' and 'rejection of discrimination' were purposely positively formulated to create alignment with the positive formulation of questions belonging to other FDVs, the constructs themselves do hold what can be considered a 'negative' (actively rejecting) component that may have mitigated the correlations with other FDVs. However, a second-order factor structure, in which 'rejection of intolerance' and 'rejection of discrimination' served as indicators of a separate latent construct as opposed to the remaining FDVs, was not the best fit for our data and indicated a discrepancy between the observed and model-implied data.



Associations of FDVs with other citizenship constructs underscore the validity of the instrument. Most FDVs had moderate to strong correlations with related constructs like civic knowledge, attitude, and self-estimated skills. The FDVs 'understanding of others' and 'freedom of speech' stood out somewhat because they were less related to civic knowledge. This indicates that students who highly underscore these FDVs do not necessarily obtain high scores on civic knowledge and, vice versa, that students who obtain high scores on civic knowledge do not necessarily highly underscore 'understanding of others' and 'freedom of speech'. In this case, 'understanding of others' particularly stands out because its low correlation with civic knowledge contrasts its high correlation with civic attitude and skills. A possible explanation may be that the other FDVs are more content-related to fundamental rights and obligations of democratic citizens, whereas 'understanding of others' entails more of a 'prosocial' component that not necessarily requires civic knowledge and, moreover, seems strongly related to a democratic attitude and self-estimated civic skills. In addition, an explanation for the low correlation of both 'understanding others' and 'freedom of speech' with civic knowledge may be that the development of civic knowledge follows a different pattern than, for example, civic attitude. Whereas knowledge generally builds up, attitudes can shift in either direction. Nonetheless, this finding contradicts studies that show that knowledge and corresponding attitudes are positively related (Lauglo, 2013). It could be that correlations between FDVs and civic knowledge become stronger when the sample consists of older (secondary school) students who have gained a more solid basis of civic knowledge – like in the study of Lauglo (2013).

Future studies may want to compare the support for democratic values across different groups. Hence, measurement invariance was assessed to determine what parameters can be meaningfully compared across groups. The results indicated that for all FDVs, at least (partial) metric invariance could be established. This allows for comparison of associations between the FDVs and the underlying items across sex, home language, school type (RPE versus SNPE),

grade (5 versus 6) and migration background. For migration background, scalar invariance could be established, which additionally allows for comparison of mean scores of the FDVs across students with and without a migration background. In that respect, it could be concluded that young students with a migration background underscored the values of 'equality', 'freedom of speech', 'tolerance', 'rejection of intolerance', and 'rejection of discrimination' significantly less compared to students without a migration background. For the other FDVs, there was no difference between students with and without migration backgrounds. As we have no information available on the country of origin, such as whether this entails equally or less democratic countries, caution is needed when interpreting these results. In addition, even when such information is available, it is not necessarily the case that individuals migrating from less democratic countries underscore FDVs to a less extent. In fact, the opposite may be true and may have even been a reason for migration in the first place.

While for all FDVs, the minimum required level of (partial) metric invariance was established (cf. Little, 2013), the measurement instrument can be further developed to establish (partial) scalar invariance across sex, home language, school type, and grade to also allow for the comparisons of latent means across these groups. This can be done by exploratory zooming in on answer patterns, mean scores, factor loadings, and item-rest correlations of questionnaire items to detect whether certain questionnaire items are hindering the establishment of a higher level of measurement invariance and require improvement or replacement. Meanwhile, the FDVQ can be used to gain information on interrelations between FDVs (i.e., students who score high on 'autonomy' also obtain high scores on 'tolerance'), to compare the associations between latent constructs and the underlying questions across groups (i.e., the relationship between the first question and 'equality' is stronger for boys than for girls), and to interpret the latent mean scores for a group as a whole (i.e., on average, students in primary education obtained a mean score of 3.61 on 'sense of responsibility').

The research findings also pointed out that students in primary education generally highly support FDVs, with the highest support for 'rejection of discrimination' and the lowest support for 'understanding others'. The high support for FDVs also manifested in other studies examining democratic values in students (Janmaat, 2018; Munniksma et al., 2017). A study in the British context, for example, found that students of approximately 23 years old supported the values democracy, individual liberty, the rule of law, and respect for people of different backgrounds and religions to a great extent (Janmaat, 2018). In addition, a longitudinal study conducted in secondary education in the Netherlands indicated that students in grade 8 (approximately 14 years old) highly supported FDVs like freedom, equality and solidarity –

like they did in grade 7 in a previous measurement (Van der Meer et al., 2021). Concerning the generally high support for FDVs, it should be noted that the current study took place in the Netherlands, which is, according to the Democracy Index[8], considered a 'full democracy' since the index started in 2006 (Economist Intelligence Unit, 2022). Underscoring FDVs to a high extent may be more obvious among students living in a country considered a 'full democracy' compared to students living in countries marked as a 'flawed democracy', a 'hybrid regime' or an 'authoritarian regime'.



Despite the overall high support for FDVs, our findings also shed light on the manifestation of disagreeing responses. To illustrate, in this study, the FDV 'rejection of intolerance' had the highest percentage of students who 'totally disagreed', indicating that when asked about the inclusion and acceptance of students with a different opinion, belief, behaviour or world view, relatively more students did not underscore the FDV as compared to the other FDVs. Previous research also pointed out that, despite high support in absolute terms, Dutch secondary school students underscored equal rights across sex, migration background and ethnicity to a lesser extent than in surrounding countries (Munniksma et al., 2017). The manifestation of more disagreeing responses to the FDV of 'rejection of intolerance' in this study, as well as the differences in support between the Netherlands and surrounding countries in previous research, both underscore the importance of monitoring support for FDVs among students periodically (or even longitudinally) to detect possible changes in support – in line with the presumption that democracy is not a fixed state of government but needs to be actively maintained. The latter is also underscored by De Winter (2004), who poses that in a democracy, there is always a lingering risk of negligence and disinterest of citizens or even unvarnished attacks by those who seek to force their totalitarian value system onto everyone.

It is important to note that when asking students to indicate their attitude towards FDVs, no information is gathered on situations in which FDVs are applied in practice. Such situations may require students to choose between conflicting FDVs. In future research, the measurement can be elaborated by asking students to choose between contradictory FDVs, for example, by providing a dilemma that reflects the tension between 'freedom of speech' and 'rejection of discrimination' or between 'autonomy' and 'sense of responsibility'. Such dilemmas were also used in research by Vaessen et al. (2022), who discussed three democratic issues (i.e., majority rule and minority rights, freedom of speech and non-discrimination, conditional political

[8] The Democracy Index is an initiative of the Economist Intelligence Unit that ranks countries based on five categories (i.e., the electoral process and pluralism, functioning of government, political participation, political culture, and civic liberties) (Economist Intelligence Unit, 2022).

representation and equal political representation) with adolescents in vocational education. By qualitatively investigating FDVs in relation to each other, the findings provide useful insight into students' consideration of democratic dilemmas when applied to everyday practice – in addition to quantitatively asking students to indicate to what extent they underscore FDVs.

This study examined whether it was possible to measure young students' attitude towards FDVs in a reliable and valid way. It can be concluded that the FDVQ is a reliable and valid instrument that can be used to measure and monitor students' attitude towards eight FDVs – which are widely conceived to be important for a pluriform and democratic society to persevere. Furthermore, the findings demonstrated that young students underscored FDVs to a great extent. These insights are relevant for research and educational practice and contribute to improving the quality and effectiveness of citizenship education, which prepares today's youth to make a positive contribution to a sustainable and resilient future democracy.

# Chapter 5
# Improving Citizenship Competences: Towards an Output-driven Approach in Citizenship Education*

**Abstract**

Scholars are increasingly paying attention to the characteristics of effective citizenship education. The systematic use of data to maximise student learning, also called an output-driven approach, is often presented as a powerful predictor of student outcomes. However, its effectiveness has not been studied in citizenship education. Therefore, this paper aims to theoretically reflect on whether an output-driven approach is also feasible for citizenship education.

We distinguish five building blocks of an output-driven approach and elaborate on their applicability in citizenship education. While doing so, we draw attention to the normative notion in citizenship education and the quality and availability of measurement instruments for citizenship competences. Both may challenge the application of an output-driven approach, particularly given the relatively young tradition of measuring citizenship competences.

We conclude that an output-driven approach in citizenship education seems feasible, provided that the characteristics of citizenship education are carefully considered.

*Keywords:* Citizenship education; Measuring citizenship competences; Output-driven approach, Data-use in education; Social outcomes of education.



*This chapter is based on: Hoek, L., Munniksma, A., & Dijkstra, A. B. (2022). Improving citizenship competences: Towards an output-driven approach in citizenship education. *Journal of Social Science Education, 21*(3), 23-43. https://doi.org/10.11576/jsse-4303

## Introduction

Increased diversity in society, continuing individualisation, growing polarisation and hardening in the social climate (Mattei & Broeks, 2018; Organisation for Economic Co-operation and Development, 2010) underscore the need for social cohesion (Dijkstra, 2012). To meet this need, many countries are stressing the importance of strengthening citizenship in society through citizenship education (Eurydice, 2017; Schulz et al., 2018). In this approach, citizenship education is important in countering threats to tolerance, equality, individual freedom and social cohesion, and democracy in general, such as lack of trust in democratic processes (Eurydice, 2017). It does so by fostering mutual respect, promoting fundamental democratic values, and strengthening students' social and civic competences (Ten Dam et al., 2011).

Following up on the growing emphasis on citizenship education, scholars are increasingly paying attention to the characteristics of effective citizenship education and to gaining insight into its outcomes in terms of citizenship competences (Cleaver et al., 2006; Daas et al., 2016; Ireland et al., 2006; Schulz et al., 2018). The results in this field point to the importance of providing students with an open classroom climate that fosters a respectful exchange of diverse opinions and enables students to practice posing an opinion independently (Geboers et al., 2013; Isac et al., 2014; Maurissen, 2020; Wanders, Dijkstra, et al., 2020). In addition, creating opportunities for students to learn and practice democracy (Isac et al., 2014) and engage in political activities in school (Hoskins et al., 2017) make a difference in students' civic competencies. However, research on the effectiveness of schools in promoting citizenship is still scarce (Coopmans et al., 2020; Dijkstra, De la Motte, & Eilard, 2014).

Scholars studying school effectiveness in other educational domains, such as reading and mathematics, have frequently presented the systematic use of data as a potentially powerful predictor of learning outcomes (Van Geel et al., 2016; Van Kuijk et al., 2016). The systematic use of data to maximise student learning is often referred to as an output-driven approach (Visscher & Ehren, 2011). An output-driven approach can help teachers make informed educational decisions (Kippers et al., 2018; Staman et al., 2017) and base these decisions on more than intuition, experience and gut feelings (Ingram et al., 2004). Throughout the years, various scholars have demonstrated the positive effects of an output-driven approach in the domain of academic achievement.

For example, Van Geel et al. (2016) found that a two-year data-based decision-making intervention in schools improved student achievement in terms of student arithmetic scores. In addition, extensive training of primary school administrators in interpreting and using student

learning analytics strengthened arithmetic achievement among students in the US (Carlson et al., 2011). Using a pretest-posttest control group design, Van Kuijk et al. (2016) found that setting standards and performance goals for students, applying formative assessment and data use, and acquiring relevant instructional skills and content and curriculum knowledge improved students' reading achievement in primary education. In addition, Slavin, Cheung, Holmes, Madden, and Chamberlain (2013) followed primary and middle schools in adopting a data-driven reform intervention designed to help school leaders understand student data and change teaching and learning based on these data. The findings show direct positive effects of the intervention on students' reading and arithmetic achievement. Similar results were found by Förster, Kawohl, and Souvignier (2018), who conducted their research in primary schools in Germany. Students whose reading achievement was frequently monitored by teachers showed higher reading fluency than students in a control group. This moderate effect appeared to be stable for at least two years. Such long-term effects were also demonstrated in the context of New Zealand (Lai et al., 2009). The authors followed students longitudinally for three years. The findings show positive effects on reading comprehension of ethnic minority students with low socioeconomic backgrounds.



Whereas most studies concluded that an output-driven approach in education effectively improves teaching and its outcomes, in some studies, no clear-cut evidence was found to support this conclusion. Comparing Dutch primary schools in an experimental group (using a digital formative assessment tool) with a control group (using their regular spelling instruction and materials), Faber and Visscher (2018) found that the use of the digital formative assessment tool did not affect third-grade student spelling achievement. According to the authors, possible explanations for this finding are the aesthetics of the tool and the number of assignments the students completed. Staman et al. (2017) found that a two-year data-based decision-making training course for teachers, principals, and academic coaches in primary schools in the Netherlands did not improve student arithmetic achievement compared to a control group. However, they found interaction effects indicating more positive results for students with lower pre-test scores and students with low SES backgrounds, which is in line with Van Geel et al. (2016). In an attempt to explain the lack of direct effects, Staman et al. (2017) mentioned the duration of the intervention (which might have been too short) or the implementation of the intervention (which might have been too fragmented). On a central condition, they also mentioned the poor data literacy of teachers. During the intervention, it appeared that the teachers experienced difficulties in translating the results of the data analysis to more specific teaching strategies.

Al in all, there seems to be substantial empirical evidence suggesting positive effects of an output-driven approach on student outcomes in the domain of academic achievement. Yet, despite the demonstrated effectiveness of output-driven teaching, such an approach has not been extensively applied to citizenship education (Eurydice, 2017). This is a remarkable observation when we consider that, next to preparation for the labour market, identity formation and social and cultural upbringing are widely conceived as the main goals of education (Dijkstra, De la Motte, & Eilard, 2014). These goals align with the qualification (i.e., the domain of education that prepares students for the labour market) and socialisation (i.e., the domain of individual and social development fostering individual potential and participation in society) function of education. The importance of research on data use in both the qualification and socialisation domain of education is gradually being recognised (Dijkstra et al., 2018; Mandinach & Jimerson, 2016; Schildkamp, 2019).

**The Present Study**

One of the crucial elements of the successful development of students' social and civic competences is a focus on learning outcomes (Dijkstra, De la Motte, & Eilard, 2014; European Commission, 2009; Eurydice, 2017). In the same way student outcome data in subjects like reading and arithmetic provide information on learning outcomes, student outcome data in citizenship education provide insight into students' civic knowledge, attitudes, and skills. In addition, student outcome data in citizenship education can guide and optimise students' learning processes, adjust learning contents to students' needs, and construct the curriculum. In this respect, monitoring citizenship competences gives schools a better understanding of the effects of their teaching, helps them adjust their curricula to student needs and helps to improve the quality of teaching and learning (Dijkstra, De la Motte, Ehren, et al., 2014). Monitoring citizenship competences also contributes to a broader and better understanding of the quality of schools and prevents the narrowing down of education, which may be an unintended consequence of evaluations focusing solely on the qualification function (cf. Koretz, 2008).

However, the applicability of an output-driven approach in citizenship education is not as self-evident as it is in a subject like mathematics. This may be due to certain characteristics specific to citizenship education, such as that it is usually offered in projects and themes rather than in a separate subject, that measurement in this domain is relatively new, and that measurement instruments are not yet widely available to schools. Schools may also be reluctant towards measurement in citizenship education because of its normative notion: the question of what are considered 'sufficient' citizenship competences is not as evident as the similar question for mathematics competences (i.e., where everybody agrees that 1 + 1 = 2). However,

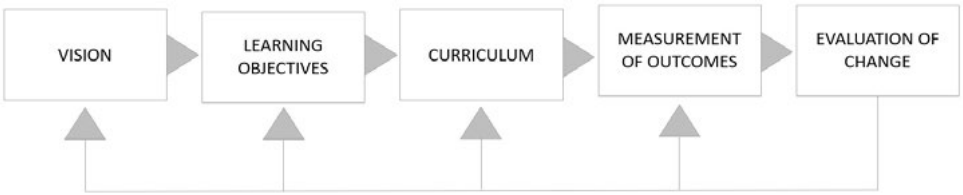
in line with the potential of an output-driven approach as demonstrated with subjects like reading and mathematics, it is interesting to examine whether an output-driven approach would also be applicable in citizenship education. This rationale leads to the central question in the present study: To what extent can an output-driven approach be applied to citizenship education? We seek to answer this question by presenting the building blocks of an output-driven approach and examining their applicability to the context of citizenship education.

## Building Blocks of an Output-driven Approach

Based on theoretical and empirical insights on the effectiveness of an output-driven approach, five main building blocks of an output-driven approach can be distinguished (Figure 1). The first one concerns the formulation of a vision, in which the school thinks of the overarching goal of (a specific domain of) their education. The second building block concerns these learning objectives, usually embodied by more detailed goals in which all main components (differentiated into knowledge, attitudes, and skills) become visible. This helps to translate the schools' vision into precise learning content. The third concerns the design of the school's curriculum so that the learning objectives can be met. The fourth building block involves the measurement of the outcomes. At last, the school evaluates whether the desired learning effect has occurred and what follow-up steps seem appropriate. Such an output-driven approach is generally conceived as a cyclic process (Kippers et al., 2018; Ledoux et al., 2009). This section elaborates on these building blocks of an output-driven approach and provides hands-on examples of how it could be applied to citizenship education.

**Figure 1**

*Model of the Output-driven Citizenship Education Approach*




### Vision

The point of departure in an output-driven approach in education is the formulation of an overarching *vision* or purpose. A vision indicates for which purpose a school educates its students: it is a broad indication of the contents to be covered and the learning goals to be achieved (Eurydice, 2017). A vision is important because it enables schools to structure their curriculum and set learning objectives according to this vision and because it enables clear communication about the contents of citizenship education with teachers, students and parents

(Hilbers et al., 2010). It is essential that the vision is shared by all school professionals (Schildkamp & Poortman, 2015), also when a vision dynamically changes throughout the years. An example of a vision for citizenship education is: *Students are prepared to be part of, and participate in, a pluriform and democratic society and, as a means to this, are tolerant and respectful towards differences in terms of culture, ethnicity, religion, sexual orientation and physical appearance; and underscore democratic values and constitutional rights.*

**Learning Objectives**

After having formulated a vision, the school should set clear *learning objectives*. Whereas a vision is generally formulated in broader terms, learning objectives are more specific goals that help a school realise its vision (Van der Kleij et al., 2015). Learning objectives must be worded precisely (Ledoux et al., 2009) and, similar to the vision, shared by all school professionals (Schildkamp & Poortman, 2015). Learning objectives may focus on the desired student achievement level or on identifying gaps between the current and desired achievement levels. In line with the previous illustration, an example of a learning objective for citizenship education is: *Students can recognise a situation where a person is being discriminated against.*

**Curriculum**

Once a school has developed a vision and accompanying learning objectives, the school designs the *curriculum* in a way that the learning objectives can be realised. According to Leeman, Nieveen, Beer, and Steen (2020), who built on the work of Goodlad, Klein, and Tye (1979), schools can focus on three aspects in this respect: (1) the essence, such as the vision, goals, content and pedagogical approaches (i.e., the substantive perspective); (2) the processes that are needed to design, develop, implement and improve the curriculum (i.e., the professional design perspective); and (3) the political processes and social interactions within the school that influence the curriculum design (i.e., the socio-political perspective). Following our learning objective above, an example of curriculum design for citizenship education is: *In line with legal guidelines and the school-specific vision and learning objectives (i.e., substantive perspective), the school designs the curriculum in a way that students learn what is, and what is not considered discrimination; and how non-discrimination is secured in the constitutional law – including why this is important for a democratic society. While doing so, the school aligns the goals, teacher and student activities and assessment (i.e., professional design perspective). In addition, the school strives for a democratic school culture in which school professionals are involved in the design and decision-making activities related to the school's general educational aims (i.e., socio-cultural perspective).*

**Measurement of Outcomes**

In the fourth step of an output-driven approach, schools should *collect data* on the extent to which the learning objectives are achieved (Kippers et al., 2018; Van der Kleij et al., 2015). Here, it is vital that schools use reliable and precise information about individual students or the school to secure the quality of the data (Schildkamp & Poortman, 2015; Van Geel & Keuning, 2016; Visscher & Ehren, 2011). In this respect, using standardised tests strengthens the robustness of the data because the measurement is similar for all students (Van Geel & Keuning, 2016; Visscher & Ehren, 2011). Standardisation can be achieved by using a fixed, comprehensive set of assignments for all students and eventually comparing student scores to a set standard or peer scores. Schools are encouraged to use multiple data sources to enable insight into differential effects, such as (standardised) student assessment data and demographic information (Hoogland et al., 2016). In citizenship education, an example of measurement is: *Students fill in a questionnaire that provides them with cases of a justified distinction between individuals and cases of discrimination. Students are asked to use their knowledge to identify the cases as such.*



**Evaluation of Change**

The output-driven approach ends with an *evaluation of change*. In this building block, the school evaluates whether the desired change has occurred by following two steps: (a) analysing and interpreting the data and (b) determining possible follow-up steps.

First, schools should *analyse and interpret the data* that have been collected. In other words, schools give meaning to the data by deciding whether the data align with the learning objectives. In citizenship education, absolute norms to indicate proficiency levels in citizenship competence are scarce and require development. Another way to interpret and assess citizenship competences is a relative or benchmark approach: results can be compared to those of similar groups of schools, regions or the national average. This approach is often used in large-scale comparative studies (cf. Cleaver et al., 2006; Ireland et al., 2006; Schulz et al., 2018; Ten Dam et al., 2011). An example of the analysis and interpretation of data in citizenship education is: *The school interprets and analyses data compared to a similar group of schools. This comparison may point out that the students of the initial school score, on average, substantially lower on the questionnaire than students of the comparison group. By zooming in on details using student background information, the school can pinpoint that boys scored lower than girls.*

Second, schools should *determine possible follow-up steps*. Based on how a school evaluates the brought-about change, the school can decide if, and if so, what follow-up steps

need to be taken. The school can do so by looking at the curriculum (e.g., a topic of citizenship education is missing or not well developed in the current curriculum) or the learning process (e.g., the contents of the curriculum are in line with the learning objectives, but require (more) differentiation in the delivery at the student level). However, if the learning outcomes are not as expected, there is often no clear solution of what needs to be altered to change this. Thus far, there is no clear causal relationship between the characteristics of the curriculum and the learning outcomes. Other factors may influence learning outcomes, such as student characteristics or characteristics of parental upbringing. In this respect, Love (2000) states that extensive dialogue among school professionals should determine the possible follow-up steps. In this dialogue, sufficient time should be reserved, and the participants should focus on exploring possible assumptions and solutions before deciding on follow-up steps. An example of the determination of possible follow-up steps in citizenship education is: *The school adapts the curriculum so that boys and girls are more likely to benefit equally from the contents. The school does so by adding role-plays to the curriculum. Students practice by alternating roles and perspectives in matters concerning discrimination – including boys taking a female stand and vice versa. In addition, the school ensures more interaction between boys and girls in the role plays.*

**Output-driven Citizenship Education: Chances and Challenges**

Based on these building blocks, we propose an output-driven model for citizenship education (Figure 1) derived from the output-driven approach applied for academic achievement (e.g., reading and arithmetic). This model has the potential to help school professionals adjust the curriculum contents of citizenship education to what students need, optimise educational outcomes at the group or school level, and improve the quality of citizenship education. However, the differences between educational outcomes such as arithmetic achievement and reading skills on the one hand, and educational outcomes such as social and civic outcomes on the other hand, require attention to determine whether the approach is possible in citizenship education. Hence, in this section, we draw attention to characteristics of citizenship education that require careful consideration when applying the model for an output-driven approach to citizenship education. These characteristics are the normative notion of citizenship, the quality and availability of measurement instruments to measure citizenship competences, and the young tradition of measuring citizenship competences within educational practice.

**The Normative Notion of Citizenship**

The meaning of citizenship and what it entails requires theoretical and empirical attention because, at least to some extent, it is a normative (cf. Eidhof, Ten Dam, Dijkstra, & Van de Werfhorst, 2016) and, therefore, essentially contested concept (Westheimer & Kahne, 2004). In practice, the perceptions of what is considered citizenship vary across schools. Consequently, the same applies to the learning objectives of citizenship education (Dijkstra et al., 2015). In this variety of perceptions of citizenship, a distinction can be made between consensus goals and contested goals (Eidhof et al., 2016). Consensus goals can be defined as general democratic, shared goals agreed upon in a democratic society, for example, equality and non-discrimination; they are typically codified in national constitutions and reflected in human rights documents. Contested goals, such as autonomy and obedience goals, are more specific and often disputed.



The normative notion of citizenship and the variety of perspectives on citizenship may lead to a situation in which the development of a vision and the specification of detailed learning objectives is considered complex and requires choices that schools feel insufficiently equipped for or reluctant to make. This may be especially true when schools struggle to fit citizenship education into what they value as a school or where the school community is divided in this respect (for example, tensions may arise between autonomy and obedience to authority).

However, if clear learning objectives are lacking, it is not possible to measure to what extent the desired proficiency level has been achieved. In this respect, there seems to be a significant difference with, for example, reading and arithmetic, where learning objectives are usually simpler to objectify, and fewer school-specific choices are required. The absence of concrete learning objectives in citizenship education in many European countries (Eurydice, 2017) challenges an output-driven approach to citizenship education. While not neglecting this challenge, we use the remainder of this section to emphasise that the perceived normative notion of citizenship education may be more limited in scope than is often assumed and is not an insurmountable obstacle in an output-driven approach to citizenship education (Dijkstra, De la Motte, Ehren, et al., 2014).

Contrary to the idea that measuring student citizenship competences has substantial disadvantages or that it is difficult or even undesirable due to its normative nature (Biesta, 2009, 2017; Koretz, 2008; Rothstein, 2009), there is broad consensus in democratic societies about the main content of citizenship education, across different perspectives on society and in schools with varying value orientations (Van Goethem et al., 2020). This provides a relevant

point of departure for meaningful measurement of citizenship competences. The essence of this rationale is the *regula aurea* or Golden Rule (Wattles, 1996). This widely held ethic of reciprocity (i.e., to treat others as you would like others to treat you) is revealed in fundamental rights such as equality, non-violence and non-discrimination and defined in national constitutions and the Universal Declaration of Human Rights.

Such fundamental rights, together with democratic values and skills aimed at dealing with diversity and conflicting values, provide the point of departure for fundamental and universal moral teaching and can be considered a shared, central core of citizenship education (Sta. Maria, 2017; Van Goethem et al., 2020; Wattles, 1996). Although values may differ for specific themes and domains, and social tasks may be viewed differently by different people and contexts, many social tasks are broadly underwritten. As such, they enable consensus about what is morally right and socially effective in a particular situation (Schneider et al., 1996). In addition, the underlying structure of social competences and the developmental processes regarding the formation of social competences is considered robust across cultures and phases of life (Halberstadt et al., 2001). Whereas this does not mean that there are no contested concepts and value differentiation in citizenship education, it does mean that a set of essential, broadly shared values is available in citizenship education. This can serve as a point of departure for formulating a vision and accompanying learning objectives.

Thus, although values of the school take a more prominent role in citizenship education than in, for example, reading and arithmetic, this does not necessarily lead to subjective assessments (Dijkstra et al., 2015). It is possible and meaningful to measure the fundamental values in citizenship education. This is underscored by large-scale international quantitative measurements in which the wide-ranging perspectives on citizenship are usually formulated in terms of knowledge, skills and attitudes. Examples include the International Civic and Citizenship Study (ICCS) (Schulz et al., 2018) and the Citizenship Competences Questionnaire (CCQ) (Ten Dam et al., 2011). However, for themes and domains where value diversity prevails, such as the contested goals of citizenship education, criteria for assessment may require school-specific norms (Dijkstra, De la Motte, Ehren, et al., 2014), and a generic measurement instrument may be less suitable. This also entails that the school team takes extra steps in deliberating the school-specific norms for these themes, which could raise an additional challenge in implementing an output-driven approach to citizenship education.

**Quality and Availability of Measurement Instruments**

To date, several measurement instruments for citizenship education have been developed. We have already noted the high-quality instruments developed within international

contexts such as ICCS (Schulz et al., 2018) and national contexts, for example, the Citizenship Education Longitudinal Study (CELS) for the UK (Cleaver et al., 2006) and the CCQ in the Netherlands (Ten Dam et al., 2011). However, reliable and valid instruments to measure citizenship competences are not widely available to schools (Daas et al., 2016; Ledoux et al., 2013; Ten Dam et al., 2011). Most high-quality measurement instruments mentioned above are primarily used for research and policy purposes and are not available for schools to assess their education.

Although measuring citizenship competences is possible, it is not yet widely performed (Eurydice, 2017). One of the explanations could be the limited instruments available to schools and the fact that schools' interest in measuring citizenship competences is relatively new. The latter seems related to the perception that measuring social outcomes in education is complex or even undesired due to its normative notion. However, complete restraint from measurement in citizenship education may not be necessary because schools can monitor the results of citizenship competences independently (Peschar et al., 2010) and without absolute standards of what level of competence is considered 'sufficient'.

In fact, measurements of citizenship competences often use relative comparison (Cleaver et al., 2006; Ireland et al., 2006; Schulz et al., 2018; Ten Dam et al., 2011). Such benchmark-based approaches amount to interpreting the measurement outcomes in terms of the distribution of outcomes in a relevant comparison group, for example, by comparing groups of students and outcome differences between schools or countries. In addition, comparing results over time may also be helpful. In a benchmark-based approach, the point of departure is not so much an absolute standard based on the desired levels of proficiency but an evaluation based on the results achieved by other schools (or student groups or countries). Although this is not so much an indication of whether the level achieved is 'sufficient', it does explain how these results deviate from the learning process outcomes in other – comparable – schools. Therefore, a relative standard shows if and how outcomes lag or compare favourably with expected results based on the group mean (and its distribution).

Thus, on the one hand, a relative standardisation based on a comparison over time or with a comparison group can be beneficial and, depending on the chosen benchmark, provide a better understanding of how much growth has been realised and how much future growth may be possible while considering school and student characteristics. However, on the other hand, the number of valid and reliable measurement instruments available to schools has been scarce for more than a decade. While this powerfully underscores the need to develop

measurement instruments that can be used for school improvement, it also poses a significant challenge in implementing an output-driven approach to citizenship education.

**The Short Tradition of Measuring Citizenship Competences**

Although socialisation is a central responsibility of education, many countries have only recently begun measurements in citizenship education (Eurydice, 2017; Organisation for Economic Co-operation and Development, 2007). More than a third of European countries do not specify national rules or recommendations for suitable classroom assessment methods where citizenship education is concerned (Eurydice, 2017), which implies that this needs to be specified at the school level. However, there is no generally accepted assessment practice. This means that schools to which this applies often only have a vague idea of the results of their citizenship education. Consequently, there is a limited basis for designing and developing a curriculum geared to the student's situation, often not based on valid and reliable data on student performance.

In addition to the challenges encountered when formulating a vision, the short tradition of measuring citizenship competences also comes with challenges of data literacy. School administrators, teachers, and others involved in the analysis of data on citizenship competences should possess data literacy, i.e., "the ability to transform information into actionable instructional knowledge and practices by collecting, analysing, and interpreting all types of data (assessment, school climate, behavioural, snapshot, longitudinal, moment-to-moment, etc.) to help determine instructional steps" (Mandinach & Gummer, 2016, p. 14). Data literacy comprises a combination of an understanding of data and standards, methodological knowledge and practices, curricular knowledge, pedagogical content knowledge and an understanding of how children learn (Mandinach & Gummer, 2016) as well as ethical and responsible use of data (e.g., on how to ensure student privacy) (Gummer & Mandinach, 2015).

Challenges of data literacy do not limit themselves to citizenship education: they are also known in measurements of, for example, reading and arithmetic achievement. However, data literacy requires even more attention in citizenship education because thinking in terms of civic outcomes is still relatively new to many schools. Moreover, knowledge of measuring instruments is often lacking, and the dominant notion is still that citizenship education does not lend itself to results measurement. Sometimes, this notion is accompanied by the view that measuring citizenship competences is undesirable. Therefore, in addition to the general attention required for data literacy as a general condition for realising an output-driven approach, attention must also be paid to the schools' unfamiliarity with the possibilities and value of standardised assessment.

Thus, the caution and reluctance on the part of school professionals towards implementing new 'boosters of technology' in education (Shirley, 2011, p. 194), such as the use of data, is not limited to citizenship education but is also common in other parts of the curriculum. However, in reading and arithmetic, school professionals have a long tradition of measuring and monitoring student learning outcomes, for example, through student monitoring system software (Kamphuis & Moelands, 2005; Ledoux et al., 2009; Verhaeghe et al., 2015; Visscher & Ehren, 2011). Consequently, school professionals may be more used to analysing and interpreting student data in this part of the curriculum. For citizenship education, the relatively young tradition of measuring student citizenship competences may reinforce the caution about analysing and interpreting student data. This forms a challenge to implementing an output-driven approach to citizenship education. However, because most primary schools are familiar with data-driven approaches to measuring students' social and socio-emotional competences for a long time (cf. Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011), it may also be possible to move beyond the feeling of caution about the measurement of citizenship competences.



**Discussion**

Whereas evaluating and improving education plays a crucial role in subjects that belong to the qualification function of education, it is less common in subjects that belong to the socialising function of education. This certainly applies to citizenship education, which has grown in importance in recent years. In fact, promoting active citizenship in education has become "one of the main objectives for education systems throughout Europe" (Eurydice, 2017, p. 7). In many cases, citizenship education can be characterised by a content-driven approach rather than an output-driven approach in which measurement outcomes are considered an essential aspect of good education.

This study theoretically reflected whether an output-driven approach could be fruitful and possible in citizenship education. Using literature on output-driven approaches in subjects like reading and arithmetic, we proposed an output-driven model for citizenship education. The five building blocks of the resulting citizenship education output-driven model are (1) the formulation of a vision; (2) setting clear learning objectives; (3) designing and implementing the curriculum according to these learning objectives; (4) measurement of the outcomes; and (5) evaluating whether the previous steps have led to the desired change. The last step consists of (5a) analysis and interpretation of the data and (5b) determination of possible necessary follow-up steps. We conclude that the output-driven model provides a good point of departure for constructing an output-driven approach for citizenship education, provided that sufficient

attention is paid to the distinctive and challenging characteristics of citizenship education. Several comments can be made in this context, the most important of which are summarised below.

Firstly, citizenship education consists of a normative notion, meaning that the vision and learning objectives (i.e., building blocks 1 and 2 in the output-driven model) may differ over schools and contexts. This marks an essential difference between citizenship education and education in reading and arithmetic since, in the latter, goals and levels of achievement are usually seen as relatively 'objective' and as the outcome of mainly functional, often academically oriented, considerations, which enable objective deliberation in the event of disagreements. Moreover, in the case of reading and arithmetic, detailed guidelines based on formal regulations or scientific knowledge are often available for determining the content and outcomes to be achieved. Currently, this is less true for the school's socialisation function. This is regularly perceived as an obstacle to measuring citizenship competences (i.e., step 4 in the output-driven model) and often leads to a reluctant attitude towards measuring in the socialisation domain of education. While not neglecting the normative notion of citizenship education, we believe it is also important to emphasise the broad consensus underlying the central goals and content of citizenship education. Most of the wide variety of content, methods, and interpersonal contacts in everyday citizenship education adhere to the principle of 'do to others as you would have them do to you' and fundamental democratic values and human rights, leading to a set of shared fundamental values. These shared fundamental values provide a solid base for the generic measurement of citizenship competences in an output-driven approach. However, an output-driven approach is more challenging for school-specific or contested goals (Eidhof et al., 2016) because it requires schools to think about what these values are, how the learning outcomes can be evaluated, and what is considered a 'sufficient' result.

Secondly, suitable measurement instruments for citizenship education are still being developed. Also, the schools' access to these instruments may differ between countries (i.e., step 4 in the output-driven model). Previous studies (Daas et al., 2016; Dijkstra, 2015; Peschar et al., 2010; Ten Dam et al., 2011), as well as large-scale international studies such as ICCS, CCQ and CELS, have shown that measuring citizenship competences is possible and meaningful if sufficient attention is paid to validity, reliability, and measurement aims. In the ideal situation, various standardised instruments are available for schools to measure student citizenship competences as part of an output-driven approach. The availability of such an instrument, however, is not a necessary condition for applying the output-driven model in

citizenship education. While measurement instruments are being developed, schools can be assisted in adopting and using more low-key, standardised and non-standardised methods (such as teacher evaluations) and qualitative methods (such as rubrics, cf. Daas et al., 2016). Instruments that are not standardised or awaiting standardisation, instruments constructed by the school or even teacher evaluations using set criteria and learning goals can also facilitate the evaluation of change in the output-driven model. Although a data-informed procedure is an important feature of the output-driven approach, a quality improvement effect can also be expected from the gradual switch from a content-driven approach to an output-driven approach. Working with concrete learning objectives and evaluating to what extent these have been achieved is essential as a first step. While developing standardised measurement instruments for citizenship competences in the long run, schools may use such alternative methods as temporary substitutes.



Regardless of the measurement instrument a school uses, understanding how learning outcomes in citizenship education can and cannot be interpreted is essential. For example, to date, there is too little known about how citizenship competences of students develop over the years. This means that it is not possible to track the learning outcomes of the same students over the years and interpret these in terms of progression. However, it is possible to compare students' learning outcomes to those of a relevant comparison group in each subsequent measurement. In this respect, previous studies have pointed to a 'dip' in citizenship competences of students in the lower grades of secondary education (Cleaver et al., 2006; Geboers et al., 2015; Keating et al., 2010). While this decrease may have various underlying reasons, it is most informative to compare the citizenship competences of students in this exact grade to a comparison group because it accounts for the 'dip' in citizenship competences that is common for all students in this grade. The caution needed to interpret the learning outcomes rightfully forms an important challenge to an output-driven approach in citizenship education, not least because it requires school professionals to be familiar with how the results can and cannot be interpreted.

Thirdly, citizenship education is usually offered in projects and themes throughout the curriculum in many schools rather than in a separate course. As a result, many school professionals share responsibility for the quality of citizenship education (i.e., step 3 of the output-driven model). As is the case in many situations of shared responsibility, this leads to the question of who bears responsibility for evaluating and interpreting the findings and their translation into curriculum development (i.e., step 4 in the output-driven model). Organising the measurement, data analysis, and interpretation of student citizenship competences is one of

the necessary steps. It requires explicit organisation, assigned responsibility and attention from the school management (Schildkamp & Poortman, 2015), certainly in a culture where thinking in terms of results and measuring outcomes is not self-evident. Moreover, it requires that those responsible for the quality of citizenship education understand how the measurement works, how the findings can be interpreted, know the measurement's limitations, and draw reasoned inferences about student achievement (Streifer, 2004). Ideally, this takes place in a school environment that ensures a positive focus on student measurement and conveys that the primary purpose of an output-driven approach is to contribute to continuous school improvement (Lachat, 2001). In sum, an output-driven approach to citizenship education requires specific skills and devotion from school professionals and can therefore be challenging and demanding. Some common educational challenges that are not unique to citizenship education, such as dealing with data literacy, challenges in data analysis and interpretation and ways to deal with reluctance towards implementing new technological tools in education, may be especially demanding in the case of citizenship education, which does not have a specific place in the curriculum and where responsibilities are not demarcated.

Although we will not discuss this in more detail here, an output-driven approach may also contribute to educational accountability and provide a better understanding of the extent to which the intended goals are achieved (Ehren & Dijkstra, 2014). Generally speaking, educational accountability is important since education uses public funds to reach shared goals. This is even more relevant to citizenship education. Here, at its minimum, accountability focuses on understanding the results of education in the form of the students' social and citizenship competences and may also include the quality of the educational process and the curriculum content. However, because of the social relevance of citizenship education and its value-bound nature, public dialogue and accountability concerning goals and results are, too, eminently crucial in this area and can contribute to improving its quality (Ehren & Dijkstra, 2014). Although educational accountability is not the primary goal of an output-driven approach, stakeholders such as parents, the local community, and the government may be interested in the extent to which the aims and goals in citizenship education are achieved and how changes concerning the initial situation are evaluated. Moreover, the nature of citizenship education entails that commonly shared goals for citizenship education should be openly accounted for to enable public deliberation of precisely those goals. This contributes to transparency and provides information to those involved in citizenship education: students, their parents, policymakers, and society. This does not alter the fact that the primary user and addressee of the outcomes of an output-driven approach are schools, with the teachers and

school management as the main target group, certainly when it comes to student and class level results. Considerations of accountability, as far as relevant, therefore primarily relate to the results at the school level.

To conclude, this study explored the applicability of an output-driven approach in citizenship education. The theoretical reflection in this study concludes that such an approach to citizenship education is possible and worthwhile and offers a fruitful step toward strengthening citizenship education and insights into student outcomes. The model of an output-driven approach presented in this study and the theoretical reflection of its challenges for citizenship education provides a basis for further development of the approach and its testing in educational practice. Together with developing and implementing the output-driven model in citizenship education, identifying its effects on the development of student citizenship competences and the development of the school curriculum in citizenship education ask for attention. An important limitation of the present study is that it primarily focused on how the school can improve citizenship education. This makes the approach more suitable for countries where citizenship education is decentralised (i.e., the school determines what is considered 'sufficient' in citizenship competences). It makes it less ideal for countries where citizenship education is centralised mainly (i.e., schools need to conform to national norms). Nevertheless, a challenging road lies ahead: although an output-driven approach in the curriculum part focusing on qualification is increasingly being acknowledged in the literature and receives recognition in the field, there still seems to be a distance between expressed interests in output-driven approaches and the actual use of data in the context of school improvement (Geijsel et al., 2010), and this certainly applies to citizenship education, where more barriers must be overcome. However, our theoretical reflection of challenges and chances for using an output-driven approach in citizenship education shows a good point of departure.

# Chapter 6

## General Discussion and Conclusion

Most European countries have formally placed the task of teaching students the competences needed to participate in a pluriform and democratic society with schools via citizenship education (Eurydice, 2017). Despite emerging research in this field, there is still a limited understanding of the factors that contribute to the quality of citizenship education (Coopmans et al., 2020). Interestingly, studies focusing on improving students' ability in reading and mathematics have frequently pointed to the positive effects of using learning outcomes as a basis for educational decisions (Van Geel et al., 2016; Van Kuijk et al., 2016). In the context of citizenship education, this has not been studied. While not neglecting the substantial differences between education focused on qualification (e.g., via reading and mathematics) and socialisation (e.g., via citizenship education), it was deemed worthwhile to examine the role of students' learning outcomes in improving the quality of citizenship education.



Hence, the preceding chapters examined the following overarching research question: 'To what extent are standardised measurement instruments able to provide for the psychometric quality and data allowing for insight into civic learning outcomes, and how can insight into civic learning outcomes be used to improve the quality of citizenship education in primary education?' As discussed in the first chapter, the studies in this dissertation were focused on standardised measurement instruments because they are generally more thoroughly developed and studied than unstandardised measurement instruments and offer better opportunities for comparison of the results across all participants. In addition, this research question has been delineated to the context of primary education, including special needs primary education, which is often overlooked in the literature. The studies in this dissertation focused on measurement invariance as one of the psychometric quality standards of measurement instruments (chapter 2), used a standardised measurement instrument to provide insight into the learning outcomes and characteristics of citizenship education in regular primary education and special needs primary education (chapter 3), developed and described the psychometric quality of a new standardised measurement instrument to measure young students' attitudes towards fundamental democratic values (chapter 4), and constructed a model for output-driven citizenship education to examine how insight into the civic learning outcomes of students can be used to improve the quality of citizenship education (chapter 5). In this final chapter, the main findings are summarised and interpreted.

**Summary of Main Findings**

This dissertation included four studies, of which the main findings are highlighted in this section. The first study (chapter 2) focused on the measurement invariance of a standardised measurement instrument based on a generic conceptualisation of citizenship competences. Standardised measurement instruments are often used to make cross-sectional or longitudinal inferences or to make comparisons across groups, such as studies pointing to robust differences in citizenship competences based on sex, socioeconomic position, and migration background (Dijkstra et al., 2015; Geijsel et al., 2012; Ireland et al., 2006; Kerr et al., 2007). Such cross-group comparative research is relevant to provide insight into whether students, regardless of their background characteristics, have equal opportunity to acquire citizenship competences. An important prerequisite for cross-group comparisons is the establishment of measurement invariance (Meredith, 1993). Measurement invariance indicates whether an instrument measures the same across groups (Isac et al., 2019; Steinmetz et al., 2009). Despite its importance for making cross-group inferences, measurement invariance is an aspect of the psychometric quality of standardised measurement instruments that received little scientific attention – particularly in within-country research (Steinmetz et al., 2009). In addition, as the interpretation of constructs can change over time (Putnick & Bornstein, 2016), it is important to periodically assess whether measurement instruments still comply with psychometric quality standards regarding measurement invariance. This applies particularly to citizenship education, which can be considered a dynamic construct that takes on meaning and moves along with changes in society (cf. Mattei & Broeks, 2018). Hence, this study assessed the measurement invariance of a standardised measurement instrument for citizenship competences developed over a decade ago to see whether the instrument still measures the same across groups in a more recent sample.

The findings in this chapter indicated that two-thirds of the measurement instrument allowed meaningful comparisons across sex, socioeconomic position, and migration background. In one-third of the cases, where the psychometric quality standards regarding measurement invariance could not be met, caution is placed on using these scales for cross-group comparisons. In these cases, the measurement instrument can still be used to describe all students in a group, grade or school. Moreover, exploratory analyses can highlight what questionnaire items need to be improved or replaced to develop the measurement instrument further. This chapter underscored that, given the dynamic nature of citizenship constructs, the psychometric quality of measurement instruments used to capture these constructs needs to be periodically assessed.

In the second study (chapter 3), standardised measurement instruments are used to provide insight into the characteristics and citizenship competences of students in regular primary education and special needs primary education. In line with the attention to and importance of citizenship education over the last decade, research into the learning outcomes and factors contributing to these learning outcomes has increased (Coopmans et al., 2020; Dijkstra et al., 2015; Geboers et al., 2013). For example, providing students with an open classroom climate for discussion, a safe school environment and fostering positive relationships among peers and between students and teachers were often positively linked to students' civic learning outcomes (Isac et al., 2019; Sampermans et al., 2018; Wanders, Dijkstra, et al., 2020; Wanders, Van der Veen, et al., 2020). However, such insights are predominantly limited to studies conducted in schools for regular education. In special needs primary education, information about students' citizenship competences and characteristics of citizenship education is largely missing. The exception is a descriptive study among Dutch students in special needs primary education in 2021 (cf. Slijkhuis et al., 2023), which found that students in special needs primary education generally obtained lower scores on citizenship competences than students in regular primary education. This research finding was the starting point of this study. As such, the third chapter examined whether characteristics of citizenship education and citizenship competences differ across regular primary education versus special needs primary education and whether the relationship between attending special needs primary education and citizenship competences changed under the influence of characteristics of citizenship education. For example, whether the relationship between attending special needs primary education and citizenship competences is more positive for students who experience a positive teacher-student relationship.

The findings of this study demonstrated that students in special needs primary education obtained lower scores on citizenship knowledge, self-estimated skills, and attitude than students in regular primary education. In addition, students in special needs primary education experienced the classroom climate to be more open to discussion, but also experienced more unsafety in the school environment and lower quality teacher-student and student-student relationships than students attending regular primary education. Contrary to expectations, the relationships between attending special needs primary education and citizenship knowledge, attitude, and self-estimated skills were not significantly influenced by students' experienced unsafety in the school environment, the teacher-student or student-student relationship, or students' perceived openness of classroom climate. While no probable explanation for these results seemed available, it must be noted that students in special needs primary education are

no clearly defined group (Pijl et al., 2008) and cope with diverse learning and behavioural difficulties (Van der Veen et al., 2010). For applied skills, the negative main effect of attending special needs primary education was weaker when students experienced more positive relationships with the teacher and with fellow students and when students experienced a more open discussion climate in the classroom. An exploration of alternative interpretations of the interaction effects, when attending special needs primary education is seen as the moderator, pointed out that the relationships between two characteristics of citizenship education (i.e., the teacher-student relationship and an open classroom climate) and applied skills are positive when students attend special needs primary education (and not significant when students attend regular primary education). However, in most cases, the relationship between characteristics of citizenship education and citizenship competences of students did not change based on whether students attended special needs primary education. By demonstrating that standardised measurement instruments are able to provide insight into the learning outcomes and characteristics of citizenship education in both regular and special needs primary education, this study contributed to the available scientific knowledge in this field and set out an important foundation for output-driven citizenship education.

The third study (chapter 4) described the construction of a new standardised measurement instrument. The Fundamental Democratic Values Questionnaire was constructed to capture young students' attitude towards fundamental democratic values. The underscoring of fundamental democratic values by citizens, such as equality, tolerance and non-discrimination, is important for a democracy to persevere (Dekker & Den Ridder, 2016; Levinson, 1999; Welzel & Dalton, 2014). Moreover, it contributes to the peaceful co-living of different (groups of) individuals in a pluriform society (De Winter, 2004). Schools are one of the socialising agents that carry the task of fostering such fundamental democratic values in students as part of citizenship education (Council of Europe, 2018). A standardised measurement instrument gives schools insight into students' learning outcomes and helps schools adjust the curriculum to students' needs. In addition, a challenge that schools encounter when measuring and monitoring students' citizenship competences is the normative notion, meaning that what is considered ('sufficient') competence in citizenship is normative (Eidhof et al., 2016) and, therefore, essentially contested (Westheimer & Kahne, 2004). A measurement instrument focused on fundamental democratic values meets this challenge of normativity in citizenship constructs. It does so by focusing on universal values which are fundamental for democracy and, therefore, less susceptible to debate (e.g., broad consensus can be found on the value that all individuals are of equal value). Such a standardised measurement instrument

aimed at monitoring students' attitude towards fundamental democratic values was not yet available to schools.

Substantive empirical support was found for the validity and reliability of the Fundamental Democratic Values Questionnaire. Also, most fundamental democratic values were moderate to strongly related to each other and with other citizenship constructs, which underscores the validity of the measurement instrument. Moreover, eight reliable scales could be distinguished. In terms of measurement invariance, a sufficient level could be established in all eight scales, which allows for a meaningful comparison of associations and (in one case) means across sex, migration background, home language, educational sector, and grade. Furthermore, the results of the first measurement of the Fundamental Democratic Values Questionnaire indicated that most young students underscored fundamental democratic values to a great extent. Mean scores were highest (3.70 on a four-point Likert scale) for 'the rejection of discrimination' and lowest (3.15 on a four-point Likert scale) for 'understanding of others', and disagreeing responses (i.e., students who 'totally disagree' with a statement) manifested most in the scale 'rejection of intolerance'. A comparison of means across migration background, which the analyses of measurement invariance allowed for, demonstrated that students with a migration background underscored 'equality', 'freedom of speech', 'tolerance', 'rejection of intolerance' and 'rejection of discrimination' significantly less than students without a migration background. However, as we have no additional information available, such as whether this entails students with a migration background from more, equally or less democratic countries, or students in a specific level of educational attainment or socioeconomic position, caution is preserved when interpreting these results. In sum, this chapter described the construction of a valid and reliable standardised measurement instrument that provided valuable information on young students' attitude towards fundamental democratic values and demonstrated that students agreed with fundamental democratic values to a great extent.

The fourth study (chapter 5) theoretically reflected on how insight into students' civic learning outcomes can be used to improve the quality of citizenship education. By doing so, a model for output-driven citizenship education was proposed. An output-driven approach has been well-studied in the educational domain that focuses on qualifying students for participating in future education, the job market and the economy broadly perceived, for example, by teaching reading and mathematics (Van Geel et al., 2016; Van Kuijk et al., 2016). Contrarily, in the educational domain that focuses on socialising students for participation in a pluriform and democratic society, an output-driven approach has not been studied (Eurydice, 2017). As such, the question of whether an output-driven approach could also be applied to

citizenship education has received little attention. While not neglecting the substantial differences between education in the qualification and socialisation domains, the considerable empirical evidence demonstrating the positive effects of an output-driven approach on students' learning outcomes in the qualification domain underscored the potential of such an approach for citizenship education. As such, derived from the literature, five building blocks of an output-driven approach were distinguished: (1) the formulation of a vision, which is the basis for (2) concrete learning objectives, which leads to (3) curriculum design, of which the school can (4) measure the learning outcomes, which are the basis for (5) evaluation and possible necessary follow-up steps (cf. Kippers et al., 2018; Ledoux et al., 2009; Van der Kleij et al., 2015; Van Geel et al., 2016; Van Kuijk et al., 2016).

Departing from these building blocks, this study theoretically reflects on two characteristics of citizenship education that make applying the building blocks to citizenship education less self-evident. The first characteristic is the normative notion in citizenship education, meaning that what is considered ('sufficient') competence in citizenship is normative. When clear guidelines on the desired level of student achievement are largely missing, it is expected from schools to decide what civic learning outcomes are important to foster and when these learning outcomes are at the desired level. This chapter emphasised that, amidst normativity, common ground can be found in values conceived as fundamental for democracy. In this respect, normativity may be more limited in scope than often assumed and not an insurmountable obstacle to output-driven citizenship education. A focus on the common ground may even help schools to deal with the challenges of normativity, as it provides a relatively 'objective' point for departure in deciding on the learning objectives in citizenship education.

The second characteristic of citizenship education that makes applying the building blocks to citizenship education less self-evident is the limited availability of quality measurement instruments for schools to measure and monitor the learning outcomes in citizenship education. While the construction of the Fundamental Democratic Values Questionnaire, as described in the third study, contributed to overcoming this challenge, the need for ongoing development of measurement instruments in citizenship education is advocated. Both the normativity and the availability of quality measurement instruments for schools challenge the application of an output-driven approach in citizenship education. This is particularly the case given the relatively young tradition of measuring the learning outcomes in this domain and other challenges not unique to citizenship education, such as the data literacy

of school professionals. Yet, this chapter concluded that applying an output-driven approach to citizenship education seems feasible when these characteristics are carefully considered.

In sum, the individual chapters shed light on the ability of standardised measurement instruments to provide for the psychometric quality and data allowing for insight into civic learning outcomes and examined how insight into civic learning outcomes can contribute to improving the quality of citizenship education. As such, the studies in this dissertation were focused on the psychometric quality of a widely used measurement instrument (i.e., the second chapter), on the ability of standardised measurement instruments to provide for insight into civic learning outcomes in different educational sectors (i.e., the third chapter), and by focusing on the psychometric quality of a standardised measurement instrument aimed at measuring students' attitude towards fundamental democratic values (i.e., the fourth chapter). Moreover, a model was provided on the role of insight into civic learning outcomes in improving the quality of citizenship education (i.e., the fifth chapter). Recalling the overarching research question of this dissertation, it can be concluded that the ability of standardised measurement instruments to provide for the psychometric quality and data allowing insight into civic learning outcomes is generally confirmed. In addition, it can be concluded that insight into civic learning outcomes offers an opportunity to improve the quality of citizenship education, provided that specific challenges are carefully considered. The next section further elaborates on the implications of this opportunity and challenges.



**Implications**

Using an output-driven approach to gain insight into civic learning outcomes and the extent to which learning objectives have been met offers an opportunity to improve the quality of citizenship education. As described in the introduction of this dissertation, there is a discrepancy between the desired and accomplished achievement levels in citizenship education in the Netherlands (Munniksma et al., 2017; Slijkhuis et al., 2021, 2023; Wagenaar et al., 2011). This discrepancy has been quite robust throughout the years and underscores the need to improve the quality of citizenship education in the Netherlands. In most European countries, including the Netherlands, thinking in terms of concrete learning objectives and measuring whether the curriculum meets the achievement of these learning objectives is less common practice (Eurydice, 2017). This contrasts the qualification domain of education, where students' learning outcomes are more frequently systematically monitored to indicate whether learning objectives have been met. Notwithstanding the challenges of applying an output-driven approach to citizenship education, following an output-driven approach may suit as guidance for schools to shift the attention from curriculum-centred citizenship education (e.g.,

organising a mock election or a project week on sexual diversity) to systematic and goal-oriented citizenship education (i.e., deciding on concrete learning objectives, for example, learning about democratic decision-making and hearing everyone's voice; designing a curriculum in such a way that it is expected to contribute to reaching the learning objectives, for example, with components such as a mock election, school policy that includes the voice of students, and lessons in which knowledge about how democracy works is transferred; and, at last, monitoring whether the curriculum suffices in reaching the learning objectives).

In the example mentioned above, monitoring learning outcomes via data collection within schools is used to serve instructional goals (i.e., using the data to improve the quality of classroom instructions, for example, by setting learning objectives based on classroom data) and school development goals (i.e., using the data to improve the functioning of the school, for example, via improvements in the policy, curriculum, or staff professionality) (Schildkamp, 2019). In this respect, gathering information on students' learning outcomes in citizenship education helps schools adjust education to what students need to improve the quality of teaching and learning (Dijkstra, De la Motte, Ehren, et al., 2014). As such, monitoring students' learning outcomes is considered essential for the successful development of citizenship competences among students (European Commission, 2009; Eurydice, 2017). In addition to instructional goals, monitoring learning outcomes can also serve accountability goals (i.e., using the data to inform stakeholders, for example, parents, on what happens in the school and on the extent to which learning objectives have been reached) (Karsten et al., 2010; Schildkamp, 2019). Even though the studies in this dissertation were predominantly focused on using learning outcomes to serve instructional and school development goals, it is also important to consider the role of learning outcomes in accountability goals.

Certain characteristics specific to citizenship education make applying an output-driven approach extra challenging and require careful consideration. These characteristics are foremost the normative notion in citizenship education and the availability and quality of measurement instruments, which are extra challenging in the context of a relatively young tradition of measuring the learning outcomes in citizenship education. To meet the challenge of the normative notion, it is suggested to consider shifting the debate from a focus on where citizenship education is different between schools to a focus on the similarities in citizenship education. In other words, we should talk more about the common ground in citizenship education, which is much greater in scope than often assumed (Dijkstra, De la Motte, Ehren, et al., 2014). This is underscored by large agreement among citizens on what should be taught in citizenship education (Dekker & Den Ridder, 2016; Van Goethem et al., 2020), in which the

essence seems to be the *regula aurea* or Golden Rule (Wattles, 1996). This widely held ethic of reciprocity is better known as 'to treat others as you want others to treat you' and is also reflected in influential international policy documents such as the Universal Declaration of Human Rights. As such, an emphasis on contested learning objectives does not entirely do justice to the everyday educational practice in citizenship education, in which most of the content, methods and interpersonal relationships adheres to the principle of 'to treat others as you want others to treat you' and to fundamental democratic values as reflected in policy documents such as the Universal Declaration of Human Rights. A focus on the common ground in citizenship education seems to be more in accordance with educational practice as it is, and offers an opportunity for standardised measurement of learning outcomes and accompanying benefits such as comparability of results. At the same time, it is acknowledged that contested and school-specific learning objectives are indeed more complex to capture using generic, standardised measurement instruments and requires more effort from schools to find or develop measurement instruments that can capture the learning outcomes of these learning objectives.



To meet the challenge of the availability and quality of measurement instruments, the ongoing development of quality measurement instruments made available for schools to use (instead of allowing access for research purposes) needs to be constructed and examined by researchers and facilitated by policy. In addition, and also to do justice to the normativity in citizenship education, it is important to have a wide variety of valid and reliable measurement instruments available. In this respect, standardised measurement instruments are preferable because their validity and reliability can more easily be ensured. In addition, standardised measurement instruments offer opportunities for relative benchmarking (i.e., comparing the learning outcomes of students in grade 8 of one school to those of students in grade 8 of other participating schools) – which offers a solution to deal with the absence of absolute norms of proficiency. However, schools can also work 'output-driven' with other assessment forms, such as portfolios, rubrics, game-based assessments, and vignettes (Daas et al., 2016, 2019). Other forms of assessment may provide additional insights. That is, standardised questionnaires capture the citizenship competences of students in a way that is most efficient and time-effective for schools but does not always offer insight into students' considerations – which, for example, a more qualitative means of measurement can capture better.

The challenges of normativity and the availability and quality of measurement instruments are strengthened amidst several general and domain-specific contextual factors. Most profoundly, data literacy of school professionals requires attention. Data literacy can be defined as "the ability to transform information into actionable instructional knowledge and

practices by collecting, analysing, and interpreting all types of data (assessment, school climate, behavioural, snapshot, longitudinal, moment-to-moment, etc.) to help determine instructional steps" (Mandinach & Gummer, 2016, p. 14). The challenge of data literacy is not unique to citizenship education (cf. Staman et al., 2017). However, it requires even more attention in citizenship education because of the relatively young tradition of measuring the outcomes and the shared responsibility for the quality of citizenship education by all school professionals. To meet these general and domain-specific contextual challenges, it is important to provide schools with more support and resources in, for example, interpreting learning outcomes to see whether the learning objectives have been met.

**Limitations and Future Research**

The findings in this dissertation are subject to at least two overarching limitations. First and foremost, all studies in this dissertation were conducted in the context of primary education in the Netherlands. As a result, the generalisability of the findings to other countries or other educational sectors is limited. In the Netherlands, school autonomy is deeply embedded in history and culture (Dijkstra et al., 2021). In line with the constitutional 'freedom of education', schools have the autonomy to establish a school in line with a philosophical or religious orientation and to organise the curriculum, teaching and learning in line with this orientation (De Groot et al., 2022; Dijkstra et al., 2021). The findings in this dissertation are predominantly focused on how schools can use measurement instruments to improve the quality of citizenship education under the premise that schools have the autonomy to do so. As such, the findings may be less applicable to countries with a more centralised educational system where the quality of citizenship education needs to adhere to national norms and, for example, concretising learning objectives is not considered a task for schools. However, even when national norms for citizenship education are available, the insights offered by relative benchmarking using standardised measurement instruments in an output-driven approach can still be considered meaningful. In addition, the studies in this dissertation were conducted in primary education, including special needs primary education. The findings may be generalised to the context of secondary or pre-vocational education, but relevant differences between the educational sectors need to be considered. An example is that schools for primary education are generally smaller in terms of the student population and school professionals and have teachers fixed to a class instead of to a subject, which may cause that (aspects of) an output-driven approach are easier and faster to implement. Moreover, the general (e.g., data literacy) and domain-specific (e.g., shared responsibility) challenges may be easier to overcome in a relatively smaller school team.

As a second limitation, the studies in this dissertation were predominantly focused on how measuring students' citizenship competences using standardised measurement instruments can contribute to the quality of citizenship education. However, the quality of citizenship education can also be improved by using other (non-standardised) assessment forms – but it does require more expertise, effort and time from schools, which may not be feasible for all schools. For example, to measure whether students have sufficiently learnt about democratic decision-making and hearing everyone's voice, a school can also use portfolios, classroom observations, rubrics or vignettes (cf. Daas et al., 2016, 2019) to see whether this learning objective has been met via, for example, the curriculum design that included a mock election, school policy that underscores the inclusion of students' voice, and lessons about how democracy works. Similar to standardised measurement instruments, for non-standardised measurement instruments, it is also important to meet the quality standards (chapter 2), to take into account the educational sector of students (chapter 3), and to find ways to expand the availability and quality of measurement instruments (chapter 4). Eventually, like standardised measurement instruments, non-standardised measurement instruments can also be used to improve the quality of citizenship education by using them in the fourth building block in an output-driven approach.



Based on this dissertation, no inferences can be drawn about the effectiveness of an output-driven approach as a whole or the effectiveness of individual underlying building blocks. In light of this dissertation, it was deemed important to first explore the opportunity and challenges of an output-driven approach in citizenship education and use the findings as a foundation for further research. In line with the conclusion that using an output-driven approach to gain insight into civic learning outcomes is an opportunity to improve the quality of citizenship education, this dissertation marks an important lead for future research: scholars should test the effectiveness of an output-driven approach in citizenship education regarding the learning outcomes of students.

**In Conclusion**

The studies in this dissertation each shed light on how standardised measurement instruments can provide for the psychometric quality and data allowing for insight into civic learning outcomes and examined how insight into civic learning outcomes can contribute to improving the quality of citizenship education. As such, this dissertation contributed to the emerging field of scientific knowledge on this topic. Hence, this dissertation can be summarised into four main contributions. First, this dissertation includes one of the few studies assessing the measurement invariance of measurement instruments for citizenship education in

a within-country setting and, as such, provided insight into the psychometric quality of existing measurement instruments in citizenship education. Second, this dissertation pioneered by demonstrating that standardised measurement instruments can be used to provide insight into the civic learning outcomes and characteristics of citizenship education in special needs primary education. Third, in this dissertation, a new standardised measurement instrument for fundamental democratic values was constructed. As such, this dissertation contributed to the limited availability of valid and reliable measurement instruments for citizenship education. Fourth, this dissertation includes a first and innovative exploration of the feasibility of an output-driven approach to citizenship education and provided more insight into its opportunity and challenges. In sum, this dissertation gives valuable insight into the ability of standardised measurement instruments to provide for the psychometric quality and data allowing for insight into civic learning outcomes and examined how this can be used to improve the quality of citizenship education.

## References


Adelson, J. L. (2012). Examining relationships and effects in gifted education research: An introduction to structural equation modeling. *Gifted Child Quarterly*, *56*(1), 47–55. https://doi.org/10.1177/0016986211424132

Alevriadou, A., & Lang, L. (2011). *Active citizenship and contexts of special education: Education for the inclusion of all students.* CiCe Thematic Network Project.

Alivernini, F., & Manganelli, S. (2011). Is there a relationship between openness in classroom discussion and students' knowledge in civic and citizenship education? *Procedia - Social and Behavioral Sciences*, *15*, 3441–3445. https://doi.org/10.1016/j.sbspro.2011.04.315

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.

Baker, J. A., Grant, S., & Morlock, L. (2008). The teacher-student relationship as a developmental context for children with internalizing or externalizing behavior problems. *School Psychology Quarterly*, *23*(1), 3–15. https://doi.org/10.1037/1045-3830.23.1.3

Barber, B. R. (1984). *Strong democracy: Participatory politics for a new age*. University of California Press.

Barrett, M., Byram, M., Aroni, A., Golubeva, I., Jouhanneau, C., Kumpulainen, K., Losito, B., Natsvlishvili, N., Rus, C., Styslavska, O., & Tranekj, L. (2021). *A portfolio of competences for democratic culture. Young learners version.* Council of Europe.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beane, J. A. (2013). A common core of a different sort: Putting democracy at the center of the curriculum: The values and skills associated with life in a democratic society should constitute the core of the curriculum. *Middle School Journal*, *44*(3), 6–14.

Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, *25*(5), 464–469. https://doi.org/10.1111/j.1467-842X.2001.tb00294.x

Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, *21*(1), 33–46. https://doi.org/10.1007/s11092-008-9064-9

Biesta, G. (2017). Education, measurement and the professions: Reclaiming a space for democratic professionality in education. *Educational Philosophy and Theory*, *49*(4), 312–314. https://doi.org/10.1080/00131857.2015.1084128

Bossaert, G., Colpin, H., Pijl, S. J., & Petry, K. (2013). Truly included? A literature study focusing on the social dimension of inclusion in education. *International Journal of Inclusive Education*, *17*(1), 60–79. https://doi.org/10.1080/13603116.2011.580464

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Routledge.

Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, *34*(2), 155–175. https://doi.org/10.1177/0022022102250225

Campbell, D. E. (2008). Voice in the classroom: How an open classroom climate fosters political engagement among adolescents. *Political Behavior*, *30*(4), 437–454. https://doi.org/10.1007/s11109-008-9063-z

Carlson, D., Borman, G. D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis*, *33*(3), 378–398. https://doi.org/10.3102/0162373711412765

Carter, E. W., Common, E. A., Sreckovic, M. A., Huber, H. B., Bottema-Beutel, K., Gustafson, J. R., Dykstra, J., & Hume, K. (2014). Promoting social competence and peer relationships for adolescents with autism spectrum disorders. *Remedial and Special Education*, *35*(2), 91–101. https://doi.org/10.1177/0741932513514618

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement*, *56*(10), 15–23. https://doi.org/10.1002/pfi.21727

Claes, E., Maurissen, L., & Havermans, N. (2017). Let's talk politics: Which individual and classroom compositional characteristics matter in classroom discussions? *YOUNG*, *25*(4), 18S-35S. https://doi.org/10.1177/1103308816673264

Cleaver, E., Ireland, E., Kerr, D., & Lopes, J. (2006). *Citizenship education longitudinal study second cross-sectional survey, 2004. Listening to young people: Citizenship education in England. Research report RR626*. ERIC Clearinghouse.

Conson, M., Mazzarella, E., Esposito, D., Grossi, D., Marino, N., Massagli, A., & Frolli, A. (2015). "Put myself into your place": Embodied simulation and perspective taking in autism spectrum disorders. *Autism Research*, *8*(4), 454–466. https://doi.org/10.1002/aur.1460

Coopmans, M., Ten Dam, G., Dijkstra, A. B., & Van der Veen, I. (2020). Towards a comprehensive school effectiveness model of citizenship education: An empirical analysis of secondary schools in The Netherlands. *Social Sciences*, *9*(9), 157. https://doi.org/10.3390/socsci9090157

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Council of Europe (2018). *Reference framework of competences for democratic culture. Descriptors of competences for democratic culture.* Council of Europe.

Creemers, B. P. M., Peters, T., & Reynolds, D. (2022). *School effectiveness and school improvement*. Routledge.

Daas, R. (2019). *Assessment of citizenship competences*. University of Amsterdam.

Daas, R., Dijkstra, A. B., & Karsten, S. (2019). Assessing young people's citizenship attitudes using rubrics. *Studies in Educational Evaluation*, *62*, 118–128. https://doi.org/10.1016/j.stueduc.2019.05.002

Daas, R., Ten Dam, G., & Dijkstra, A. B. (2016). Contemplating modes of assessing citizenship competences. *Studies in Educational Evaluation*, *51*, 88–95. https://doi.org/10.1016/j.stueduc.2016.10.003

De Groot, I., Daas, R., & Nieuwelink, H. (2022). Education for democratic citizenship in Dutch schools: A bumpy road. *Journal of Social Science Education*, *21*(4). https://doi.org/10.11576/jsse-5381

De Schaepmeester, L., van Braak, J., & Aesaert, K. (2022a). Social citizenship competences at the end of primary school: The role of socio-ethnic classroom diversity and teachers' citizenship beliefs and practices in the classroom climate. *Theory & Research in Social Education*, *50*(2), 297–326. https://doi.org/10.1080/00933104.2022.2051106

De Schaepmeester, L., van Braak, J., & Aesaert, K. (2022b). Social citizenship competences of primary school students: Does socio-ethnic diversity in the classroom matter? *Globalisation, Societies and Education*, *20*(3), 365–387. https://doi.org/10.1080/14767724.2021.1889360

De Sousa, M. L., Peixoto, M. M., & Cruz, S. F. (2021). The association between externalizing and internalizing problems with bullying engagement in adolescents: The mediating role of social skills. *International Journal of Environmental Research and Public Health*, *18*(19), 10444. https://doi.org/10.3390/ijerph181910444

De Waal, T. (2019). *Tussen kwetsbaarheid en weerbaarheid [Between vulnerability and resilience]*. Dutch National Committee 4-5th of May.

De Winter, M. (2004). *Opvoeding, onderwijs en jeugdbeleid in het algemeen belang. De noodzaak van een democratisch-pedagogisch offensief [Pedagogics, education, and youth policy in the public interest. The need for a democratic-pedagogical offensive]*. Wetenschappelijke Raad voor het Regeringsbeleid (WRR).

Dekker, P., & Den Ridder, J. (2016). *Gedeelde waarden en een weerbare democratie [Shared values and a resilient democracy]*. Sociaal en Cultureel Planbureau.

Department for Education (2014). *Promoting fundamental British values as part of SMSC in schools. Departmental advice for maintained schools.* Department for Education.

Devine, J., & Cohen, J. (2007). *Making your school safe: Strategies to protect children and promote learning*. Teachers College Press.

Dewey, J. (1923). *Democracy and education. An introduction to the philosophy of education.* Macmillan.

Dijkstra, A. B. (2012). *Sociale opbrengsten van onderwijs [Social outcomes of education]*. Vossiuspers UvA.

Dijkstra, A. B. (2015). *Moeilijk meetbare onderwijsresultaten op het sociale en maatschappelijke domein. Verkennende notitie in opdracht van de Onderwijsraad [Educational outcomes in the social and societal domain that are difficult to measure. Explorative note commissioned by the Education Council]*. Dutch Education Council.

Dijkstra, A. B., Daas, R., De la Motte, P. I., & Ehren, M. (2018). Inspecting school social quality: Assessing and improving school effectiveness in the social domain. *Journal of Social Science Education*, *16*(4), 75–84. https://doi.org/10.4119/UNIBI/JSSE-V16-I4-1640

Dijkstra, A. B., & De la Motte, P. I. (2014). *Social outcomes of education: The assessment of social outcomes and school improvement through school inspections*. Amsterdam University Press.

Dijkstra, A. B., De la Motte, P. I., Ehren, M., & Eilard, A. (2014). Discussion. School inspections and school improvement in the social domain. The assessment of social outcomes of education. In A. B. Dijkstra & P. I. De la Motte (Eds.), *Social outcomes of education. The assessment of social outcomes and school improvement through school inspections* (pp. 189–211). Amsterdam University Press.

Dijkstra, A. B., De la Motte, P. I., & Eilard, A. (2014). Social outcomes of education. Concept and measurement. In A. B. Dijkstra & P. I. De la Motte (Eds.), *Social outcomes of education. The assessment of social outcomes and school improvement through school inspections* (pp. 29–50). Amsterdam University Press.

Dijkstra, A. B., El Khayati, N., & Vosse, A. (2014). Evaluation of social outcomes in the Netherlands. In A. B. Dijkstra & P. I. De la Motte (Eds.), *Social outcomes of education: The assessment of social outcomes and school improvement through school inspections* (pp. 103–118). Amsterdam University Press.

Dijkstra, A. B., Geijsel, F., Ledoux, G., Van der Veen, I., & Ten Dam, G. (2015). Effects of school quality, school citizenship policy, and student body composition on the acquisition of citizenship competences in the final year of primary education. *School Effectiveness and School Improvement*, *26*(4), 524–553. https://doi.org/10.1080/09243453.2014.969282

Dijkstra, A. B., Ten Dam, G., & Munniksma, A. (2021). Inequality in citizenship competences. Citizenship education and policy in the Netherlands. In B. Malak-Minkiewicz & J. Torney-Purta (Eds.), *Influences of the IEA Civic and Citizenship Education Studies: Practice, policy, and sesearch across countries and regions*. Springer International Publishing. https://doi.org/10.1007/978-3-030-71102-3

Douma, I., De Boer, A., Minnaert, A., & Grietens, H. (2022). The I of students with ID or SEBD: A systematic literature review of the self-concept of students with ID or SEBD. *Educational Research Review*, *36*, 100449. https://doi.org/10.1016/j.edurev.2022.100449

Drisko, J. (1993). The responsibilities of schools in civic education. *The Journal of Education*, *175*(1), 105–119. https://doi.org/10.1177/0022057493175001

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of

school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Economist Intelligence Unit (2022). *Democracy Index 2022 Frontline democracy and the battle for Ukraine*.

Ehren, M., & Dijkstra, A. B. (2014). Evaluation of social outcomes through school inspections. In A. B. Dijkstra & P. I. De la Motte (Eds.), *Social outcomes of education. The assessment of social outcomes and school improvement through school inspections* (pp. 51–72). Amsterdam University Press.

Eidhof, B. B. F., Ten Dam, G., Dijkstra, A. B., & Van de Werfhorst, H. G. (2016). Consensus and contested citizenship education goals in Western Europe. *Education, Citizenship and Social Justice*, *11*(2), 114–129. https://doi.org/10.1177/1746197915626084

Etzioni, A. (1996). *The new Golden Rule. Community and morality in a democratic society.* Basic Books.

European Commission (2009). *Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions. Draft 2010 joint progress report of the council and the commission on the implementation of the 'Education & Training 2010 work programme'. Key competences for a changing world*. Commission of the European Communities.

*European Convention on Human Rights*, Council of Europe (1953).

Eurostat (2021). *Migrant integration statistics*. European Commission.

Eurydice (2005). *Citizenship education at school in Europe*. Eurydice European Unit.

Eurydice (2017). *Citizenship education at school in Europe 2017: Eurydice brief*. Publications Office. https://doi.org/10.2797/818387

Faber, J. M., & Visscher, A. (2018). The effects of a digital formative assessment tool on spelling achievement: Results of a randomized experiment. *Computers & Education*, *122*, 1–8. https://doi.org/10.1016/j.compedu.2018.03.008

Farmer, T. W., Irvin, M. J., Motoca, L. M., Leung, M., Hutchins, B. C., Brooks, D. S., & Hall, C. M. (2015). Externalizing and internalizing behavior problems, peer affiliations, and bullying involvement across the transition to middle school. *Journal of Emotional and Behavioral Disorders*, *23*(1), 3–16. https://doi.org/10.1177/1063426613491286

Förster, N., Kawohl, E., & Souvignier, E. (2018). Short- and long-term effects of assessment-based differentiated reading instruction in general education on reading fluency and

reading comprehension. *Learning and Instruction*, *56*, 98–109. https://doi.org/10.1016/j.learninstruc.2018.04.009

Freire, S., Pipa, J., Aguiar, C., Vaz da Silva, F., & Moreira, S. (2020). Student–teacher closeness and conflict in students with and without special educational needs. *British Educational Research Journal*, *46*(3), 480–499. https://doi.org/10.1002/berj.3588

Frostad, P., & Pijl, S. J. (2007). Does being friendly help in making friends? The relation between the social position and social skills of pupils with special needs in mainstream education. *European Journal of Special Needs Education*, *22*(1), 15–30. https://doi.org/10.1080/08856250601082224

Galston, W. A. (2001). Political knowledge, political engagement, and civic education. *Annual Review of Political Science*, *4*(1), 217–234. https://doi.org/10.1146/annurev.polisci.4.1.217

Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, *2*, 66–70.

Geboers, E., Geijsel, F., Admiraal, W., Jorgensen, T., & Ten Dam, G. (2015). Citizenship development of adolescents during the lower grades of secondary education. *Journal of Adolescence*, *45*(1), 89–97. https://doi.org/10.1016/j.adolescence.2015.08.017

Geboers, E., Geijsel, F., Admiraal, W., & Ten Dam, G. (2013). Review of the effects of citizenship education. *Educational Research Review*, *9*, 158–173. https://doi.org/10.1016/j.edurev.2012.02.001

Geijsel, F., Krüger, M., & Sleegers, P. J. C. (2010). Data feedback for school improvement: The role of researchers and school leaders. *The Australian Educational Researcher*, *37*(2), 59–75. https://doi.org/10.1007/BF03216922

Geijsel, F., Ledoux, G., Reumerman, R., & Ten Dam, G. (2012). Citizenship in young people's daily lives: Differences in citizenship competences of adolescents in the Netherlands. *Journal of Youth Studies*, *15*(6), 711–729. https://doi.org/10.1080/13676261.2012.671932

Gietz, C., & McIntosh, K. (2014). Relations between student perceptions of their school environment and academic achievement. *Canadian Journal of School Psychology*, *29*(3), 161–176. https://doi.org/10.1177/0829573514540415

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. https://doi.org/10.1007/s10964-013-0084-5

Goodlad, J. I., Klein, M. F., & Tye, K. A. (1979). The domains of curriculum and their study. In J. I. Goodlad (Ed.), *Curriculum inquiry: The study of curriculum practice* (pp. 43–76). McGraw-Hill.

Gummer, E., & Mandinach, E. (2015). Building a conceptual framework for data literacy. *Teachers College Record: The Voice of Scholarship in Education*, *117*(4), 1–22. https://doi.org/10.1177/016146811511700401

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (7th ed.).* Pearson Educational International.

Halberstadt, A. G., Denham, S. A., & Dunsmore, J. C. (2001). Affective social competence. *Social Development*, *10*(1), 79–119. https://doi.org/10.1111/1467-9507.00150

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301

He, J., & Van de Vijver, F. J. R. (2013). Methodological issues in cross-cultural studies in educational psychology. In G. A. D. Liem & A. B. I. Bernardo (Eds.), *Advancing cross-cultural perspectives on educational psychology: A festschrift for Dennis McInerney* (pp. 39–56). Information Age Publishing.

Hilbers, G., Dekkers, H., & Dijkstra, A. B. (2010). De ontwikkeling van burgerschapsonderwijs op scholen voor primair en voortgezet onderwijs [The development of citizenship education in primary and secondary education]. In J. Peschar, H. Hooghoff, A. B. Dijkstra, & G. Ten Dam (Eds.), *Scholen voor burgerschap. Naar een kennisbasis voor burgerschapsonderwijs [Schools for citizenship. Towards a knowledge base for citizenship education]* (pp. 43–60). Garant.

Hoek, L., Munniksma, A., & Dijkstra, A. B. (2022). Improving citizenship competences: Towards an output-driven approach in citizenship education. *Journal of Social Science Education*, *21*(3). https://doi.org/10.11576/JSSE-4303

Hoogland, I., Schildkamp, K., Van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B., & Dijkstra, A. M. (2016). Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. *Teaching and Teacher Education*, *60*, 377–386. https://doi.org/10.1016/j.tate.2016.07.012

Horn, J. L., & Mcardle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144. https://doi.org/10.1080/03610739208253916

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. https://doi.org/10.1080/10705519909540118

Hurwitz, S., Garman-McClaine, B., & Carlock, K. (2022). Special education for students with autism during the COVID-19 pandemic: "Each day brings new challenges". *Autism*, *26*(4), 889–899. https://doi.org/10.1177/13623613211035935

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, *106*(6), 1258–1287. https://doi.org/10.1111/j.1467-9620.2004.00379.x

Inspectorate of Education (2012). *De staat van het onderwijs 2010/2011 [The state of education 2010/2011]*. Ministry of Education, Culture and Science.

Inspectorate of Education (2014). *De staat van het onderwijs 2012/2013 [The state of education 2012/2013]*. Ministry of Education, Culture and Science.

Inspectorate of Education (2016). *Burgerschap op school. Een beschrijving van burgerschapsonderwijs en de maatschappelijke stage [Citizenship at school. A description of citizenship education and the societal internship]*. Ministry of Education, Culture and Science.

Inspectorate of Education (2019). *De staat van het onderwijs 2019 [The state of education 2019]*.

Inspectorate of Education (2022a). *De staat van het onderwijs 2022 [The state of education 2022]*. Ministry of Education, Culture and Science.

Inspectorate of Education (2022b). *Peil.Burgerschap einde basisonderwijs [The state of citizenship at the end of primary education]*. Ministry of Education, Culture and Science.

Inspectorate of Education (2023). *Peil.Burgerschap einde speciaal (basis)onderwijs [The state of citizenship at the end of special needs primary education]*. Ministry of Education, Culture and Science.

Ireland, E., Kerr, D., Lopes, J., Nelson, J., & Cleaver, E. (2006). *Active citizenship and young people: Opportunities, experiences and challenges in and beyond school. Citizenship*

*education longitudinal study: Fourth annual report (DfES Research Report 732)*. Department for Education and Skills.

Isac, M. M., Maslowski, R., Creemers, B., & Van der Werf, G. (2014). The contribution of schooling to secondary-school students' citizenship outcomes across countries. *School Effectiveness and School Improvement*, *25*(1), 29–63. https://doi.org/10.1080/09243453.2012.751035

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

Isac, M. M., Palmerio, L., & Van der Werf, M. P. C. (2019). Indicators of (in)tolerance toward immigrants among European youth: An assessment of measurement invariance in ICCS 2016. *Large-Scale Assessments in Education*, *7*(1), 1–21. https://doi.org/10.1186/s40536-019-0074-5

Janmaat, J. G. (2011). Ability grouping, segregation and civic competences among adolescents. *International Sociology*, *26*(4), 455–482. https://doi.org/10.1177/0268580910393044

Janmaat, J. G. (2018). Educational influences on young people's support for fundamental British values. *British Educational Research Journal*, *44*(2), 251–273. https://doi.org/10.1002/berj.3327

Jennissen, R., Engbersen, G., Bokhorst, M., & Bovens, M. (2018). *De nieuwe verscheidenheid. Toenemende diversiteit naar herkomst in Nederland [The new variety. Increasing diversity by origin in The Netherlands]*. Wetenschappelijke Raad voor het Regeringsbeleid (WRR).

Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion & Parties*, *15*(2), 237–264. https://doi.org/10.1080/13689880500178849

Jorgensen, T. D. (2017). Applying permutation tests and multivariate modification indices to configurally invariant models that need respecification. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01455

Jorgensen, T. D. (2021). *semTools: Useful tools for structural equation modeling*. https://cran.r-project.org/web/packages/semTools/index.html

Kamphuis, F., & Moelands, F. (2005). A student monitoring system. *Educational Measurement: Issues and Practice*, *19*(4), 28–30. https://doi.org/10.1111/j.1745-3992.2000.tb00046.x

Karsten, S., Visscher, A. J., Bert Dijkstra, A., & Veenstra, R. (2010). Towards standards for the publication of performance indicators in the public sector: The case of schools. *Public Administration*, *88*(1), 90–112. https://doi.org/10.1111/j.1467-9299.2010.01812.x

Keating, A., Kerr, D., Benton, T., Mundy, E., & Lopes, J. (2010). *Citizenship education in England 2001-2010: Young people's practices and prospects for the future. Research Report DFE-RR059.* Department for Children, Schools and Families.

Kerr, D., Lopes, J., Nelson, J., White, K., Cleaver, E., & Benton, T. (2007). *VISION versus PRAGMATISM: Citizenship in the secondary school curriculum in England : citizenship education longitudinal study : fifth annual report*. Department for Education and Skills.

Kippers, W., Poortman, C., Schildkamp, K., & Visscher, A. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*, *56*, 21–31. https://doi.org/10.1016/j.stueduc.2017.11.001

Kite, B. A., Jorgensen, T. D., & Chen, P. (2018). Random permutation testing applied to measurement invariance testing with ordered-categorical indicators. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 573–587. https://doi.org/10.1080/10705511.2017.1421467

Kline, R. B. (2015). *Principles and practices of structural equation modelling*. The Guilford Press.

Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*, *18*(1), 1–17. https://doi.org/10.1187/cbe.18-04-0064

Knowles, R. T., Torney-Purta, J., & Barber, C. (2018). Enhancing citizenship learning with international comparative research: Analyses of IEA civic education datasets. *Citizenship Teaching & Learning*, *13*(1), 7–30. https://doi.org/10.1386/ctl.13.1.7_1

Koh, K., & Zumbo, B. (2008). Multi-group confirmatory factor analysis for testing measurement invariance in mixed item format data. *Journal of Modern Applied Statistical Methods*, *7*(2). https://doi.org/10.22237/jmasm/1225512660

Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Harvard University Press.

Kuang, X., Kennedy, K. J., & Mok, M. M. C. (2018). Creating democratic class rooms in asian contexts: The influences of individual and school level factors on open classroom climate. *Journal of Social Science Education*, *17*(1), 29–40. https://doi.org/10.4119/jsse-863

Kymlicka, W. (1989). *Liberalism, community and culture*. Clarendon Press.

Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal*, *36*(5), 807–830. https://doi.org/10.1080/01411920903165603

Lachat, M. A. (2001). *Data-driven high school reform: The breaking ranks model*. LAB at Brown University.

Lai, M. K., McNaughton, S., Amituanai-Toloa, M., Turner, R., & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly*, *44*(1), 30–56. https://doi.org/10.1598/RRQ.44.1.2

Lander, V. (2016). Introduction to fundamental British values. *Journal of Education for Teaching*, *42*(3), 274–279. https://doi.org/10.1080/02607476.2016.1184459

Lauglo, J. (2013). Do more knowledgeable adolescents have more rationally based civic attitudes? Analysis of 38 countries. *Educational Psychology*, *33*(3), 262–282. https://doi.org/10.1080/01443410.2013.772773

Ledoux, G., Blok, H., Boogaard, M., & Kruger, M. (2009). *Opbrengstgericht werken: Over de waarde van meetgestuurd onderwijs [To work results-oriented: About the value of data-driven education]*. Kohnstamm Instituut.

Ledoux, G., Meijer, J., Van der Veen, I., & Breetvelt, I. (2013). *Meetinstrumenten voor sociale competenties, metacognitie en advanced skills [Measurement instruments for social competence, meta cognition and advanced skills]*. Kohnstamm Instituut.

Ledoux, G., Roeleveld, J., Van Langen, A., & Smeeets, E. (2012). *Cool Speciaal. Inhoudelijk rapport [Cool Speciaal. Substantive report]*. Kohnstamm Instituut.

Ledwell, M., & King, V. (2015). Bullying and internalizing problems: Gender differences and the buffering role of parental communication. *Journal of Family Issues*, *36*(5), 543–566. https://doi.org/10.1177/0192513X13491410

Leeman, Y., Nieveen, N., Beer, F., & Steen, J. (2020). Teachers as curriculum-makers: The case of citizenship education in Dutch schools. *The Curriculum Journal*, *31*(3), 495–516. https://doi.org/10.1002/curj.21

Levinson, M. (1999). Liberalism, pluralism, and political education: Paradox or paradigm? *Oxford Review of Education*, *25*(1–2), 39–58. https://doi.org/10.1080/030549899104116

Little, T. D. (2013). *Longitudinal structural equation modelling*. The Guilford Press.

Love, N. (2000). *Using data – getting results. Collaborative inquiry for school-based mathematics and science reform*. The Regional Alliance for Mathematics and Science Education Reform.

Mandinach, E., & Gummer, E. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. Teachers College Press.

Mandinach, E., & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. *Teaching and Teacher Education*, *60*, 452–457. https://doi.org/10.1016/j.tate.2016.07.009

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

Mansfield, E. R., & Helms, B. P. (1982). Detecting multicollinearity. *The American Statistician*, *36*(3), 158–160. https://doi.org/10.2307/2683167

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

Maslowski, R., Van der Werf, M. P. C., Oonk, G. H., Naayer, H. M., & Isac, M. M. (2012). *Burgerschapscompetenties van leerlingen in de onderbouw van het voortgezet onderwijs. Eindrapport van de International Civic and Citizenship Education Study (ICCS) in Nederland [Citizenship competences of students in the lower grades of secondary education. Final report of the International Civic and Citizenship Education Study (ICCS) in the Netherlands]*. GION.

Mattei, P., & Broeks, M. (2018). From multiculturalism to civic integration: Citizenship education and integration policies in the Netherlands and England since the 2000s. *Ethnicities*, *18*(1), 23–42. https://doi.org/10.1177/1468796816676845

Maurissen, L. (2018). *Civic engagement of adolescents*. Catholic University of Leuven.

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

Maurissen, L., Claes, E., & Barber, C. (2018). Deliberation in citizenship education: How the school context contributes to the development of an open classroom climate. *Social Psychology of Education*, *21*(4), 951–972. https://doi.org/10.1007/s11218-018-9449-7

Meade, A. W. (2005). *Sample size and tests of measurement invariance*. The 20th annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, *7*(4), 361–388. https://doi.org/10.1177/1094428104268027

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Meuleman, B., Żółtak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D. L., Billiet, J., & Schmidt, P. (2022). Why measurement invariance is important in comparative research. A response to Welzel et al. (2021). *Sociological Methods & Research*, 00491241221091755. https://doi.org/10.1177/00491241221091755

Miller, D. I. (2000). *Citizenship and national identity.* Polity.

Mishna, F. (2003). Learning disabilities and bullying: Double jeopardy. *Journal of Learning Disabilities*, *36*(4), 336–347. https://doi.org/10.1177/00222194030360040501

Munck, I., Barber, C., & Torney-Purta, J. (2018). Measurement invariance in comparing attitudes toward immigrants among youth ccross Europe in 1999 and 2009: The alignment method applied to IEA CIVED and ICCS. *Sociological Methods & Research*, *47*(4), 687–728. https://doi.org/10.1177/0049124117729691

Munniksma, A., Dijkstra, A. B., Van der Veen, I., Ledoux, G., Van de Werfhorst, H. G., & Ten Dam, G. (2017). *Burgerschap in het voortgezet onderwijs: Nederland in vergelijkend perspectief [Citizenship in secondary education: The Netherlands in comparative perspective]*. Amsterdam University Press.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*(4), 557–585. https://doi.org/10.1007/BF02296397

O'Kearney, R., Salmon, K., Liwag, M., Fortune, C.-A., & Dawel, A. (2017). Emotional abilities in children with oppositional defiant disorder (ODD): Impairments in perspective-taking and understanding mixed emotions are associated with high callous–unemotional traits. *Child Psychiatry & Human Development*, *48*(2), 346–357. https://doi.org/10.1007/s10578-016-0645-4

Organisation for Economic Co-operation and Development (2007). *Understanding the social outcomes of learning*. OECD Publications.

Organisation for Economic Co-operation and Development (2010). *Improving health and social cohesion through education*. OECD Publications.

Organisation for Economic Co-operation and Development (2017). *Understanding the socioeconomic divide in Europe*. OECD Publications.

Persson, M. (2015). Classroom climate and political learning: Findings from a Swedish panel study and comparative data. *Political Psychology*, *36*(5), 587–601. https://doi.org/10.1111/pops.12179

Peschar, J., Hooghoff, H., Dijkstra, A. B., & Ten Dam, G. (2010). *Scholen voor burgerschap. Naar een kennisbasis voor burgerschapsonderwijs [Schools for citizenship. Towards a knowledge base for citizenship education]*. Garant.

Pijl, S. J., Frostad, P., & Flem, A. (2008). The social position of pupils with special needs in regular schools. *Scandinavian Journal of Educational Research*, *52*(4), 387–405. https://doi.org/10.1080/00313830802184558

Pinto, C., Baines, E., & Bakopoulou, I. (2019). The peer relations of pupils with special educational needs in mainstream primary schools: The importance of meaningful contact and interaction with peers. *British Journal of Educational Psychology*, *89*(4), 818–837. https://doi.org/10.1111/bjep.12262

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reeskens, T., & Hooghe, M. (2010). Beyond the civic–ethnic dichotomy: Investigating the structure of citizenship concepts across thirty-three countries. *Nations and Nationalism*, *16*(4), 579–597. https://doi.org/10.1111/j.1469-8129.2010.00446.x

Reiter, S., & Schalock, R. L. (2008). Applying the concept of quality of life to Israeli special education programs: A national curriculum for enhanced autonomy in students with special needs. *International Journal of Rehabilitation Research*, *31*(1), 13–21. https://doi.org/10.1097/MRR.0b013e3282f45201

Rinnooy Kan, W. F., März, V., Volman, M., & Dijkstra, A. B. (2021). Learning from, through and about differences: A multiple case study on schools as practice grounds for citizenship. *Social Sciences*, *10*(6), 200. https://doi.org/10.3390/socsci10060200

Roorda, D. L., Zee, M., Bosman, R. J., & Koomen, H. M. Y. (2021). Student–teacher relationships and school engagement: Comparing boys from special education for autism spectrum disorders and regular education. *Journal of Applied Developmental Psychology*, *74*, 101277. https://doi.org/10.1016/j.appdev.2021.101277

Rose, C. A., & Espelage, D. L. (2012). Risk and protective factors associated with the bullying involvement of students with emotional and behavioral disorders. *Behavioral Disorders*, *37*(3), 133–148. https://doi.org/10.1177/019874291203700302

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rothstein, R. (2009). What's wrong with accountability by the numbers? *American Educator*, *33*(1), 20–33.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. Institute of Education, University of London.

Sampermans, D., Isac, M. M., & Claes, E. (2018). Can schools engage students? Multiple perspectives, multidimensional school climate research in England and Ireland. *Journal of Social Science Education*, *17*(1), 13–28. https://doi.org/10.4119/jsse-857

Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, *61*(3), 257–273. https://doi.org/10.1080/00131881.2019.1625716

Schildkamp, K., & Poortman, C. (2015). Factors influencing the functioning of data teams. *Teachers College Record: The Voice of Scholarship in Education*, *117*(4), 1–42. https://doi.org/10.1177/016146811511700403

Schneider, R. J., Ackerman, P. L., & Kanfer, R. (1996). To "act wisely in human relations:" Exploring the dimensions of social competence. *Personality and Individual Differences*, *21*(4), 469–481. https://doi.org/10.1016/0191-8869(96)00084-0

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*. International Association for the Evaluation of Educational Achievement.

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 International Publishing. https://doi.org/10.1007/978-3-319-73963-2

Schulz, W., Carstens, R., Losito, B., & Fraillon, J. (2016). *ICCS 2016 Technical report. IEA International Civic and Citizenship Education Study 2016*. International Association for the Evaluation of Educational Achievement.

Schwab, S. (2019). Friendship stability among students with and without special educational needs. *Educational Studies*, *45*(3), 390–401. https://doi.org/10.1080/03055698.2018.1509774

Schwab, S. (2021). Inclusive and special education in Europe. In U. Sharma & S. J. Salend (Eds.), *The Oxford Encyclopedia of Inclusive and Special Education* (pp. 807–819). Oxford University Press. https://doi.org/10.1093/acrefore/9780190264093.013.1230

Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement*, *79*(2), 310–334. https://doi.org/10.1177/0013164418783530

Shirley, D. (2011). The Fourth Way of technology and change. *Journal of Educational Change*, *12*(2), 187–209. https://doi.org/10.1007/s10833-011-9164-z

Slavin, R. E., Cheung, A., Holmes, G. C., Madden, N. A., & Chamberlain, A. (2013). Effects of a data-driven district reform model on state assessment outcomes. *American Educational Research Journal*, *50*(2), 371–396. https://doi.org/10.3102/0002831212466909

Slijkhuis, E., Dijkstra, A. B., Ritzema, L., De Boer, H., Naayer, H., Van der Veen, I., Stronkhorst, E., Walet, L., Kuijpers, R., Hemker, B., & Lampe, T. (2021). *Burgerschap in het basisonderwijs. Technisch rapport Peil.Burgerschap 2020 [Citizenship in primary education. Technical report of Peil.Burgerschap 2020]*. Universiteit van Amsterdam.

Slijkhuis, E., Hoek, L., De Boer, H., Hemker, B., Van der Veen, I., Naayer, H., Stronkhorst, E., & Dijkstra, A. B. (2023). *Burgerschap in het speciaal(basis)onderwijs. Technisch rapport Peil.Burgerschap 2021 [Citizenship in special needs primary education. Technical report of Peil.Burgerschap 2021]*. Universiteit van Amsterdam.

Sta. Maria, J. E. (2017). *Shu* and *z hong* as the virtue of the Golden Rule: A Confucian contribution to contemporary virtue ethics. *Asian Philosophy*, *27*(2), 100–111. https://doi.org/10.1080/09552367.2017.1318530

Staman, L., Timmermans, A. C., & Visscher, A. (2017). Effects of a data-based decision making intervention on student achievement. *Studies in Educational Evaluation*, *55*, 58–67. https://doi.org/10.1016/j.stueduc.2017.07.002

Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78–90. https://doi.org/10.1086/209528

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity*, *43*(4), 599–616. https://doi.org/10.1007/s11135-007-9143-x

Streifer, P. A. (2004). *Tools and techniques for effective data-driven decision making*. ScarecrowEducation.

Ten Dam, G., Geijsel, F., Ledoux, G., & Meijer, J. (2013). Citizenship of students and social desirability: Living apart together? *International Journal of Educational Research*, *62*, 229–238. https://doi.org/10.1016/j.ijer.2013.10.001

Ten Dam, G., Geijsel, F., Reumerman, R., & Ledoux, G. (2011). Measuring young people's citizenship competences. *European Journal of Education*, *46*(3), 354–372. https://doi.org/10.1111/j.1465-3435.2011.01485.x

Ten Dam, G., & Volman, M. (2007). Educating for adulthood or for citizenship: Social competence as an educational goal. *European Journal of Education*, *42*(2), 281–298. https://doi.org/10.1111/j.1465-3435.2007.00295.x

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Tomlinson, S. (2015). The empire disintegrates. *Ethnic and Racial Studies*, *38*(13), 2208–2215. https://doi.org/10.1080/01419870.2015.1058509

Tyler, M., & Iyengar, S. (2022). Learning to dislike your opponents: Political socialization in the era of polarization. *American Political Science Review*, 1–8. https://doi.org/10.1017/S000305542200048X

*United Nations Convention on the Rights of the Child*, General Assembly resolution 44/25 (1989).

*Universal Declaration of Human Rights*, Nations General Assembly (1948).

US Census Bureau (2020). *Estimates of the components of resident population change by race and Hispanic origin for the United States*. US Census Bureau.

https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html

Vaessen, A., Daas, R., & Nieuwelink, H. (2022). All things considered: The views of adolescents in vocational education on competing democratic values. *Citizenship, Social and Economics Education*, *21*(1), 3–21. https://doi.org/10.1177/20471734211050290

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). **mice**: Multivariate imputation by chained equations in *R*. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

Van der Kleij, F. M., Vermeulen, J. A., Schildkamp, K., & Eggen, T. J. H. M. (2015). Integrating data-based decision making, Assessment for Learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice*, *22*(3), 324–343. https://doi.org/10.1080/0969594X.2014.999024

Van der Meer, T., Wanders, F., Thijs, P., Mulder, L., Aizenberg, E., Ten Dam, G., & Van de Werfhorst, H. (2021). *Democratische kernwaarden in het voortgezet onderwijs Adolescentenpanel Democratische Kernwaarden en Schoolloopbanen jaar 2—2019/2020 [Fundamental democratic values in secondary education]*. University of Amsterdam.

Van der Veen, I., Smeets, E., & Derriks, M. (2010). Children with special educational needs in the Netherlands: Number, characteristics and school career. *Educational Research*, *52*(1), 15–43. https://doi.org/10.1080/00131881003588147

Van Geel, M., & Keuning, T. (2016). *Implementation and effects of a school-wide data-based decision making intervention*. https://doi.org/10.3990/1.9789036541190

Van Geel, M., Keuning, T., Visscher, A., & Fox, J. (2016). Assessing the effects of a school-wide data-based decision-making intervention on student achievement growth in primary schools. *American Educational Research Journal*, *53*(2), 360–394. https://doi.org/10.3102/0002831216637346

Van Goethem, A., Ten Dam, G., & Dijkstra, A. B. (2020). What does society want adolescents to know about civics? *Research Papers in Education*, 1–22. https://doi.org/10.1080/02671522.2020.1864767

Van Gunsteren, H. (1998). *A theory of citizenship: Organizing plurality in contemporary democracies*. Westview Press.

Van Kuijk, M. F., Deunk, M. I., Bosker, R. J., & Ritzema, E. S. (2016). Goals, data use, and instruction: The effect of a teacher professional development program on reading

achievement. *School Effectiveness and School Improvement*, *27*(2), 135–156. https://doi.org/10.1080/09243453.2015.1026268

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Verhaeghe, G., Schildkamp, K., Luyten, H., & Valcke, M. (2015). Diversity in school performance feedback systems. *School Effectiveness and School Improvement*, *26*(4), 612–638. https://doi.org/10.1080/09243453.2015.1017506

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, *30*(6), 1024–1054. https://doi.org/10.1080/01419870701599465

Vieno, A., Perkins, D. D., Smith, T. M., & Santinello, M. (2005). Democratic school climate and sense of community in school: A multilevel analysis. *American Journal of Community Psychology*, *36*(3–4), 327–341. https://doi.org/10.1007/s10464-005-8629-8

Visscher, A., & Ehren, M. (2011). *De eenvoud en complexiteit van opbrengstgericht werken [The simplicity and complexity of a results-oriented approach]*. Commissioned by Ministry of Education, Culture and Science.

Wagenaar, H., Van der Schoot, F., & Hemker, B. (2011). Balans Actief burgerschap en sociale integratie. Uitkomsten van de peiling in 2009 [Balance active citizenship and social integration. Results from a measurement in 2009]. *Cito*.

Wanders, F. H. K., Dijkstra, A. B., Maslowski, R., & Van der Veen, I. (2020). The effect of teacher-student and student-student relationships on the societal involvement of students. *Research Papers in Education*, *35*(3), 266–286. https://doi.org/10.1080/02671522.2019.1568529

Wanders, F. H. K., Van der Veen, I., Dijkstra, A. B., & Maslowski, R. (2020). The influence of teacher-student and student-student relationships on societal involvement in Dutch primary and secondary schools. *Theory & Research in Social Education*, *48*(1), 101–119. https://doi.org/10.1080/00933104.2019.1651682

Wardekker, W. L. (2001). Schools and moral education: Conformism or autonomy? *Journal of Philosophy of Education*, *35*(1), 101–114. https://doi.org/10.1111/1467-9752.00212

Wattles, J. (1996). *The Golden Rule*. Oxford University Press.

Welzel, C. (2021). Why the future is democratic. *Journal of Democracy*, *32*(2), 132–144. https://doi.org/10.1353/jod.2021.0024

Welzel, C., & Dalton, R. J. (2014). From allegiant to assertive citizens. In R. J. Dalton & C. Welzel (Eds.), *The Civic Culture Transformed* (1st ed., pp. 282–306). Cambridge University Press. https://doi.org/10.1017/CBO9781139600002.017

Wentzel, K. R. (2017). Peer relationships, motivation, and academic performance at school. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (pp. 586–603). The Guilford Press.

Westheimer, J. (2008). On the relationship between political and moral engagement. In F. K. Oser & W. Veugelers (Eds.), *Getting involved: Global citizenship development and sources of moral values* (pp. 17–29). Sense Publishers.

Westheimer, J., & Kahne, J. (2004). What kind of citizen? The politics of educating for democracy. *American Educational Research Journal*, *41*(2), 237–269. https://doi.org/10.3102/00028312041002237

*Wet van 23 juni 2021 tot wijziging van een aantal onderwijswetten in verband met verduidelijking van de burgerschapsopdracht aan scholen in het funderend onderwijs [Law of 23 June 2021 for the adaptation in several educational laws concerning the clarification of the citizenship education assignment to schools]*, Ministry of Education, Culture and Science (2021).

White, P. (1999). Political education in the early years: The place of civic virtues. *Oxford Review of Education*, *25*(1/2), 59–70. https://doi.org/10.1080/030549899104125

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

Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*, *81*(4), 1014–1045. https://doi.org/10.1007/s11336-016-9506-0

Yakut, A. D. (2021). Educators' experiences in special education institutions during the COVID-19 outbreak. *Journal of Research in Special Educational Needs*, *21*(4), 345–354. https://doi.org/10.1111/1471-3802.12533

Zee, M., & Roorda, D. L. (2018). Student–teacher relationships in elementary school: The unique role of shyness, anxiety, and emotional problems. *Learning and Individual Differences*, *67*, 156–166. https://doi.org/10.1016/j.lindif.2018.08.006

**Appendix 2.1**

*Model Fit of Configural, Threshold, Metric, and Scalar Model across Sex*

| | | $\chi^2$ | df | P | CFI | RMSEA | CI RMSEA | Anova | $\Delta\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|---|---|
| Attitude | | | | | | | | | | |
| | Configural | 296.642 | 10 | 0.000*** | 0.964 | 0.100 | [0.091; 0.110] | - | - | |
| | Metric | 304.165 | 14 | 0.000*** | 0.964 | 0.085 | [0.077; 0.094] | Configural | 7.5236 | 0.1107 |
| | Scalar | 567.460 | 18 | 0.000*** | 0.932 | 0.104 | [0.096; 0.111] | Metric | 263.30 | <0.001*** |
| Attitude – Acting in a socially desirable way | | | | | | | | | | |
| | Configural | 200.257 | 18 | 0.000*** | 0.964 | 0.060 | [0.053; 0.067] | - | - | - |
| | Threshold | 220.041 | 24 | 0.000*** | 0.961 | 0.054 | [0.047; 0.060] | Configural | 2.2175 | 0.8986 |
| | Metric | 210.176 | 29 | 0.000*** | 0.964 | 0.047 | [0.041; 0.053] | Threshold | 6.237 | 0.2838 |
| | Scalar | 268.797 | 34 | 0.000*** | 0.954 | 0.049 | [0.044; 0.055] | Metric | 45.632 | <0.001*** |
| Attitude – Dealing with conflicts | | | | | | | | | | |
| | Configural | 164.454 | 18 | 0.000*** | 0.993 | 0.054 | [0.046; 0.061] | - | - | - |
| | Threshold | 175.315 | 24 | 0.000*** | 0.993 | 0.047 | [0.041; 0.054] | Configural | 4.3649 | 0.6274 |
| | Metric | 175.180 | 29 | 0.000*** | 0.993 | 0.042 | [0.036; 0.048] | Threshold | 12.401 | 0.02968* |
| | Scalar | 183.975 | 34 | 0.000*** | 0.993 | 0.039 | [0.034; 0.045] | Metric | 12.074 | 0.03379* |
| Attitude – Dealing with differences | | | | | | | | | | |
| | Configural | 1132.638 | 18 | 0.000*** | 0.962 | 0.148 | [0.141; 0.155] | - | - | - |
| | Threshold | 1211.898 | 24 | 0.000*** | 0.959 | 0.132 | [0.126; 0.139] | Configural | 9.8652 | 0.1304 |
| | Metric | 1105.878 | 29 | 0.000*** | 0.963 | 0.115 | [0.109; 0.120] | Threshold | 4.1677 | 0.5255 |
| | Scalar | 1119.360 | 34 | 0.000*** | 0.963 | 0.106 | [0.101; 0.112] | Metric | 37.56 | <0.001*** |
| Knowledge | | | | | | | | | | |
| | Configural | 140.400 | 4 | 0.000*** | 0.971 | 0.110 | [0.094; 0.126] | - | - | - |
| | Metric | 148.345 | 7 | 0.000*** | 0.970 | 0.084 | [0.073; 0.096] | Configural | 7.9455 | 0.04715* |
| | Scalar | 302.460 | 10 | 0.000*** | 0.937 | 0.102 | [0.092; 0.112] | Metric | 154.11 | <0.001*** |
| Knowledge – Acting democratically | | | | | | | | | | |
| | Configural | 103.537 | 40 | 0.000*** | 0.981 | 0.024 | [0.018; 0.029] | - | - | - |
| | Scalar | 113.542 | 46 | 0.000*** | 0.980 | 0.023 | [0.018; 0.028] | Configural | 9.7076 | 0.1375 |
| Knowledge – Acting in a socially desirable way | | | | | | | | | | |
| | Configural | 21.512 | 18 | 0.254 | 0.997 | 0.008 | [0.000; 0.020] | - | - | - |
| | Scalar | 26.429 | 22 | 0.234 | 0.997 | 0.008 | [0.000; 0.019] | Configural | 4.6404 | 0.32622 |
| Knowledge – Dealing with conflicts | | | | | | | | | | |
| | Configural | 226.743 | 28 | 0.000*** | 0.943 | 0.050 | [0.044; 0.056] | - | - | - |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Scalar | 232.386 | 33 | 0.000*** | 0.943 | 0.046 | [0.041; 0.052] | Configural | 13.260 | 0.02106* |
| Knowledge – Dealing with differences | | | | | | | | | | |
| | Configural | 99.869 | 18 | 0.000*** | 0.955 | 0.040 | [0.033; 0.048] | - | - | - |
| | Scalar | 115.965 | 22 | 0.000*** | 0.948 | 0.039 | [0.032; 0.046] | Configural | 13.0913 | 0.01084* |
| | Partial | 105.885 | 20 | 0.000*** | 0.953 | 0.039 | [0.032; 0.046] | Configural | 4.9802 | 0.0829 |
| | Scalar | | | | | | | | | |
| Skill | | | | | | | | | | |
| | Configural | 99.441 | 4 | 0.000*** | 0.987 | 0.092 | [0.077; 0.108] | - | - | - |
| | Metric | 101.927 | 7 | 0.000*** | 0.987 | 0.069 | [0.058; 0.081] | Configural | 2.486 | 0.4778 |
| | Scalar | 247.994 | 10 | 0.000*** | 0.967 | 0.092 | [0.082; 0.102] | Metric | 146.07 | <0.001*** |
| | Partial | 102.50 | 9 | 0.000*** | 0.987 | 0.060 | [0.050; 0.071] | Metric | 0.56847 | 0.7526 |
| | Scalar | | | | | | | | | |
| Skill – Acting in a socially desirable way/Dealing with conflicts | | | | | | | | | | |
| | Configural | 413.516 | 10 | 0.000*** | 0.975 | 0.120 | [0.110; 0.130] | - | - | - |
| | Threshold | 434.607 | 15 | 0.000*** | 0.974 | 0.100 | [0.092; 0.108] | Configural | 6.2712 | 0.2807 |
| | Metric | 384.895 | 19 | 0.000*** | 0.978 | 0.083 | [0.076; 0.090] | Threshold | 1.9376 | 0.7472 |
| | Scalar | 474.821 | 23 | 0.000*** | 0.972 | 0.083 | [0.077; 0.090] | Metric | 69.79 | <0.001*** |
| Skill – Dealing with differences | | | | | | | | | | |
| | Configural | 33.351 | 4 | 0.000*** | 0.996 | 0.051 | [0.036; 0.068] | - | - | - |
| | Threshold | 38.813 | 8 | 0.000*** | 0.995 | 0.037 | [0.026; 0.049] | Configural | 3.3767 | 0.4969 |
| | Metric | 43.312 | 12 | 0.000*** | 0.995 | 0.032 | [0.022; 0.043] | Threshold | 5.0567 | 0.1677 |
| | Scalar | 44.883 | 14 | 0.000*** | 0.995 | 0.028 | [0.019; 0.037] | Metric | 2.591 | 0.4591 |

**Appendix 2.2**

*Model Fit of Configural, Threshold, Metric, and Scalar Model across Socioeconomic Position*

| | | $\chi^2$ | df | P | CFI | RMSEA | CI RMSEA | Anova | $\Delta\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|---|---|
| Attitude | | | | | | | | | | |
| | Configural | 195.933 | 10 | 0.000*** | 0.968 | 0.096 | [0.084; 0.108] | - | - | - |
| | Metric | 198.077 | 14 | 0.000*** | 0.968 | 0.081 | [0.071; 0.091] | Configural | 2.1444 | 0.7092 |
| | Scalar | 225.477 | 18 | 0.000*** | 0.964 | 0.075 | [0.067; 0.084] | Metric | 27.399 | <0.001*** |
| | Partial Scalar | 198.083 | 15 | 0.000*** | 0.968 | 0.078 | [0.068; 0.087] | Metric | 0.006 | 0.9381 |
| Attitude – Acting in a socially responsible manner | | | | | | | | | | |
| | Configural | 164.591 | 18 | 0.000*** | 0.962 | 0.063 | [0.055; 0.073] | - | - | - |
| | Threshold | 182.332 | 24 | 0.000*** | 0.959 | 0.057 | [0.050; 0.065] | Configural | 3.0837 | 0.7983 |
| | Metric | 194.089 | 29 | 0.000*** | 0.958 | 0.053 | [0.046; 0.060] | Threshold | 16.238 | 0.006197** |
| | Partial Metric | 173.415 | 27 | 0.000*** | 0.963 | 0.052 | [0.045; 0.059] | Threshold | 1.7855 | 0.6181 |
| | Partial Scalar | 180.591 | 32 | 0.000*** | 0.962 | 0.048 | [0.041; 0.055] | Partial Metric | 8.5067 | 0.1304 |
| Attitude – Dealing with conflicts | | | | | | | | | | |
| | Configural | 114.129 | 18 | 0.000*** | 0.993 | 0.051 | [0.043; 0.061] | - | - | - |
| | Threshold | 128.988 | 24 | 0.000*** | 0.993 | 0.047 | [0.39; 0.055] | Configural | 9.3279 | 0.156 |
| | Metric | 112.524 | 29 | 0.000*** | 0.994 | 0.038 | [0.031; 0.045] | Threshold | 1.0021 | 0.9624 |
| | Scalar | 111.648 | 34 | 0.000*** | 0.995 | 0.034 | [0.027; 0.041] | Metric | 3.5128 | 0.6215 |
| Attitude – Dealing with differences | | | | | | | | | | |
| | Configural | 740.961 | 18 | 0.000*** | 0.964 | 0.141 | [0.133; 0.150] | - | - | - |
| | Threshold | 788.920 | 24 | 0.000*** | 0.962 | 0.126 | [0.118; 0.133] | Configural | 2.3189 | 0.8882 |
| | Metric | 731.332 | 29 | 0.000*** | 0.965 | 0.110 | [0.103; 0.117] | Threshold | 8.1128 | 0.1501 |
| | Scalar | 719.695 | 34 | 0.000*** | 0.966 | 0.100 | [0.094; 0.106] | Metric | 12.266 | 0.03132* |
| Knowledge | | | | | | | | | | |
| | Configural | 109.354 | 4 | 0.000*** | 0.970 | 0.114 | [0.096; 0.133] | - | - | - |
| | Metric | 110.875 | 7 | 0.000*** | 0.970 | 0.086 | [0.072; 0.100] | Configural | 1.5212 | 0.6774 |
| | Scalar | 146.012 | 10 | 0.000*** | 0.961 | 0.082 | [0.070; 0.094] | Metric | 35.137 | <0.001*** |
| Knowledge – Acting democratically | | | | | | | | | | |
| | Configural | 68.933 | 40 | 0.003** | 0.987 | 0.019 | [0.011; 0.026] | - | - | - |
| | Scalar | 80.816 | 46 | 0.001** | 0.984 | 0.019 | [0.012; 0.026] | Configural | 10.189 | 0.1169 |
| Knowledge – Acting in a socially responsible manner | | | | | | | | | | |
| | Configural | 15.598 | 18 | 0.621 | 1.000 | 0.000 | [0.000; 0.017] | - | - | - |
| | Scalar | 19.792 | 22 | 0.596 | 1.000 | 0.000 | [0.000; 0.016] | Configural | 4.0691 | 0.3967 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Knowledge – Dealing with conflicts | | | | | | | | | | |
| | Configural | 187.003 | 28 | 0.000*** | 0.947 | 0.053 | [0.046; 0.060] | - | - | - |
| | Scalar | 180.329 | 33 | 0.000*** | 0.951 | 0.047 | [0.040; 0.054] | Configural | 1.6152 | 0.8994 |
| Knowledge – Dealing with differences | | | | | | | | | | |
| | Configural | 72.946 | 18 | 0.000*** | 0.958 | 0.039 | [0.030; 0.048] | - | - | - |
| | Scalar | 72.592 | 22 | 0.000*** | 0.961 | 0.034 | [0.025; 0.043] | Configural | 3.3881 | 0.4951 |
| Skill | | | | | | | | | | |
| | Configural | 63.587 | 4 | 0.000*** | 0.988 | 0.086 | [0.068; 0.105] | - | - | - |
| | Metric | 65.780 | 7 | 0.000*** | 0.988 | 0.064 | [0.051; 0.079] | Configural | 2.1931 | 0.5333 |
| | Scalar | 80.967 | 10 | 0.000*** | 0.986 | 0.059 | [0.048; 0.071] | Metric | 15.187 | 0.001664** |
| | Partial Scalar | 67.729 | 9 | 0.000*** | 0.989 | 0.057 | [0.045; 0.070] | Metric | 1.949 | 0.3774 |
| Skill – Acting in a socially responsible manner/Dealing with conflicts | | | | | | | | | | |
| | Configural | 263.175 | 10 | 0.000*** | 0.979 | 0.112 | [0.101; 0.124] | - | - | - |
| | Threshold | 274.545 | 15 | 0.000*** | 0.978 | 0.093 | [0.083; 0.102] | Configural | 2.1858 | 0.8229 |
| | Metric | 255.037 | 19 | 0.000*** | 0.980 | 0.079 | [0.070; 0.087] | Threshold | 7.6401 | 0.1057 |
| | Scalar | 254.560 | 23 | 0.000*** | 0.981 | 0.071 | [0.063; 0.079] | Metric | 4.1723 | 0.3832 |
| Skill – Dealing with differences | | | | | | | | | | |
| | Configural | 19.616 | 4 | 0.001** | 0.997 | 0.044 | [0.026; 0.064] | - | - | - |
| | Threshold | 22.100 | 8 | 0.005** | 0.997 | 0.030 | [0.015; 0.045] | Configural | 1.5899 | 0.8106 |
| | Metric | 25.492 | 11 | 0.008** | 0.997 | 0.026 | [0.012; 0.039] | Threshold | 3.8552 | 0.2775 |
| | Scalar | 36.466 | 14 | 0.001** | 0.996 | 0.028 | [0.017; 0.040] | Metric | 8.6192 | 0.03481* |

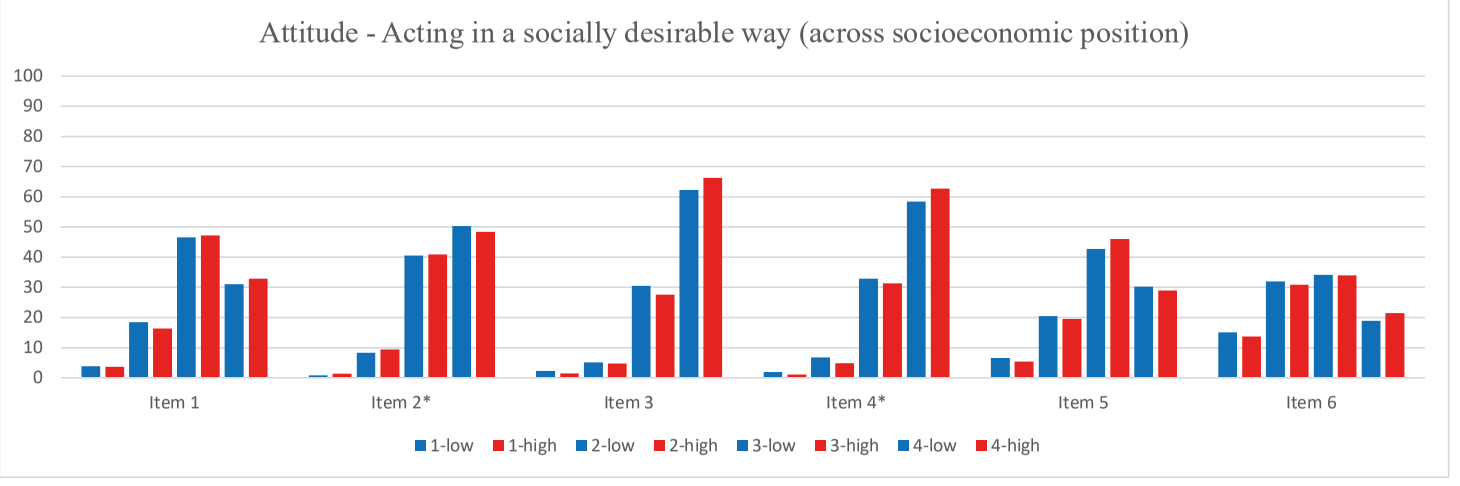
**Appendix 2.3**

*Model Fit of Configural, Threshold, Metric, and Scalar Model across Migration Background*

| | | $\chi^2$ | df | P | CFI | RMSEA | CI RMSEA | Anova | $\Delta\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|---|---|
| Attitude | | | | | | | | | | |
| | Configural | 276.123 | 10 | 0.000*** | 0.968 | 0.097 | [0.087; 0.107] | - | - | - |
| | Metric | 288.138 | 14 | 0.000*** | 0.967 | 0.083 | [0.075; 0.091] | Configural | 12.015 | 0.01724* |
| | Scalar | 554.333 | 18 | 0.000*** | 0.936 | 0.102 | [0.095; 0.110] | Metric | 266.19 | <0.001*** |
| Attitude – Acting in a socially responsible manner | | | | | | | | | | |
| | Configural | 210.179 | 18 | 0.000*** | 0.968 | 0.061 | [0.054; 0.069] | - | - | - |
| | Threshold | 230.528 | 24 | 0.000*** | 0.965 | 0.055 | [0.049; 0.062] | Configural | 4.506 | 0.6085 |
| | Metric | 224.131 | 29 | 0.000*** | 0.967 | 0.049 | [0.043; 0.055] | Threshold | 9.2647 | 0.09896 |
| | Scalar | 457.328 | 34 | 0.000*** | 0.928 | 0.066 | [0.061; 0.072] | Metric | 103.48 | <0.001*** |
| Attitude – Dealing with conflicts | | | | | | | | | | |
| | Configural | 172.091 | 18 | 0.000*** | 0.993 | 0.055 | [0.048; 0.063] | - | - | - |
| | Threshold | 181.417 | 24 | 0.000*** | 0.993 | 0.048 | [0.042; 0.055] | Configural | 5.6431 | 0.4643 |
| | Metric | 164.825 | 29 | 0.000*** | 0.994 | 0.041 | [0.035; 0.047] | Threshold | 5.4018 | 0.3688 |
| | Scalar | 185.670 | 34 | 0.000*** | 0.993 | 0.040 | [0.034; 0.045] | Metric | 14.326 | 0.01366* |
| | Partial Scalar | 170.475 | 33 | 0.000*** | 0.994 | 0.038 | [0.033; 0.044] | Metric | 7.4826 | 0.1125 |
| Attitude – Dealing with differences | | | | | | | | | | |
| | Configural | 1196.779 | 18 | 0.000*** | 0.958 | 0.152 | [0.145; 0.160] | - | - | - |
| | Threshold | 1260.470 | 24 | 0.000*** | 0.955 | 0.135 | [0.129; 0.141] | Configural | 17.618 | 0.007262** |
| | Metric | 1118.213 | 29 | 0.000*** | 0.961 | 0.115 | [0.110; 0.121] | Threshold | 8.3448 | 0.1382 |
| | Scalar | 1104.191 | 34 | 0.000*** | 0.961 | 0.105 | [0.100; 0.111] | Metric | 31.088 | <0.001*** |
| Knowledge | | | | | | | | | | |
| | Configural | 138.223 | 4 | 0.000*** | 0.972 | 0.109 | [0.094; 0.125] | - | - | - |
| | Metric | 145.666 | 7 | 0.000*** | 0.971 | 0.083 | [0.072; 0.096] | Configural | 7.4431 | 0.05904 |
| | Scalar | 176.543 | 10 | 0.000*** | 0.965 | 0.077 | [0.067; 0.087] | Metric | 30.878 | <0.001*** |

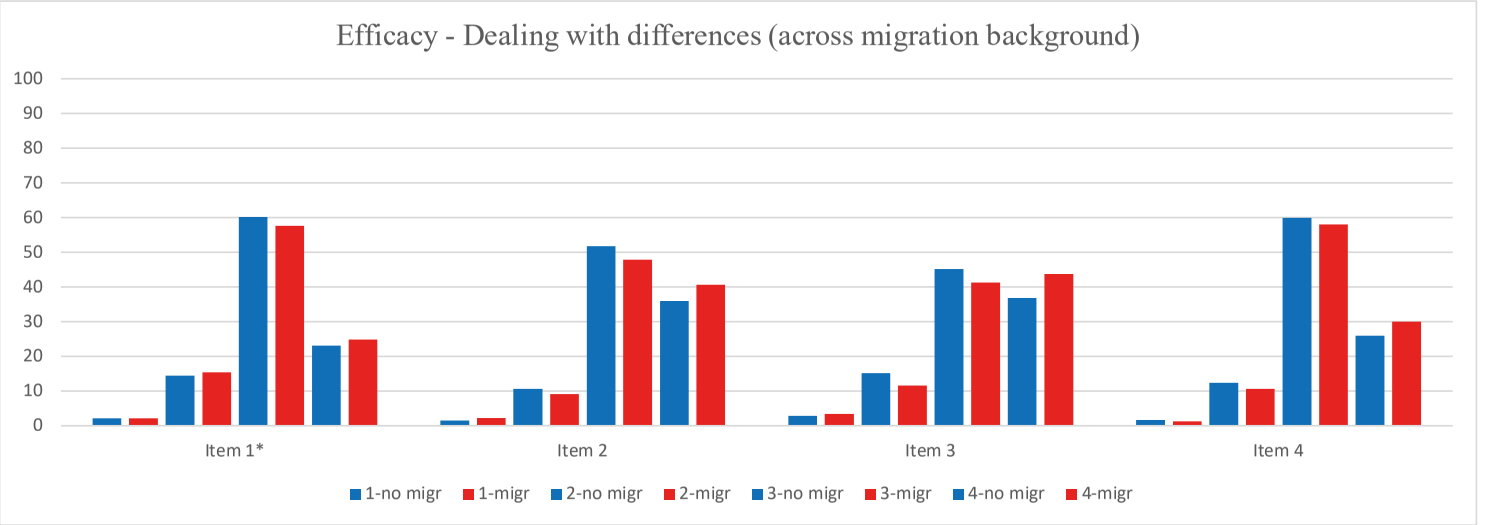
**Appendix 2.4**

*Exploratory Analyses*



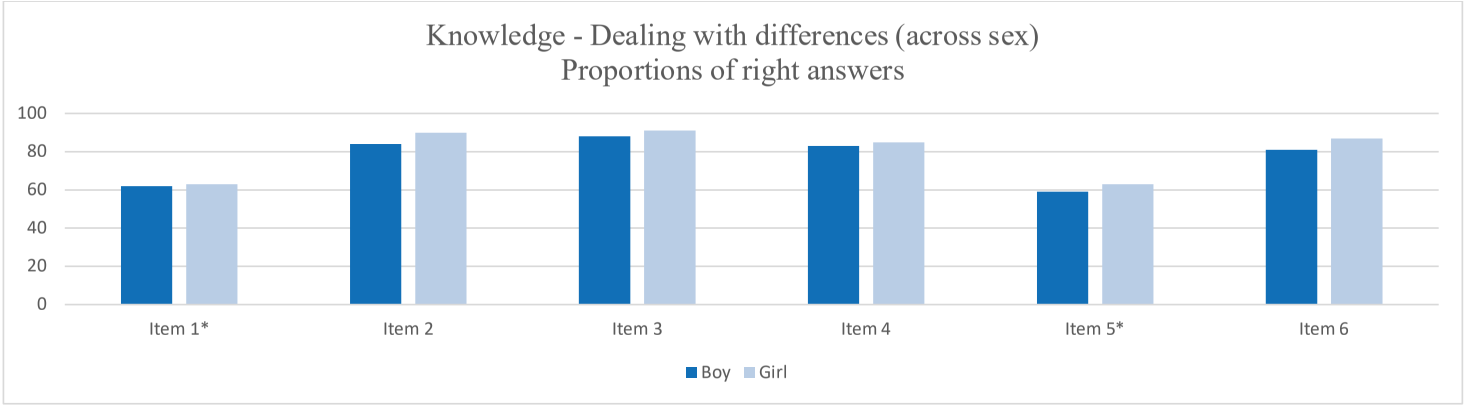

| *Attitude – Acting in a socially desirable way (across socioeconomic position)* | | Low SEP | | | High SEP | | | Difference Δ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | λ | Rir | μ | λ | Rir | Δμ | Δλ | ΔRir |
| 1 | Wealthy people taking care of less wealthy people | 3.05 | 0.486 | 0.36 | 3.09 | 0.514 | 0.36 | 0.04 | 0.028 | 0.00 |
| 2* | Standing up for a classmate when s/he is being called names | 3.41 | 0.401 | 0.29 | 3.36 | 0.494 | 0.35 | -0.05 | 0.093 | 0.06 |
| 3 | Cleaning up after having a picknick in the parc | 3.53 | 0.698 | 0.43 | 3.59 | 0.660 | 0.39 | 0.06 | -0.038 | -0.04 |
| 4* | Saying sorry if you hurt somebody | 3.48 | 0.757 | 0.48 | 3.56 | 0.656 | 0.40 | 0.08 | -0.101 | -0.08 |
| 5 | Helping in the household | 2.97 | 0.454 | 0.33 | 2.99 | 0.448 | 0.31 | 0.02 | -0.006 | -0.02 |
| 6 | Visiting a classmate who has been sick for a long time | 2.57 | 0.487 | 0.36 | 2.63 | 0.471 | 0.34 | 0.06 | -0.016 | -0.02 |

*Note.* SEP = socioeconomic position. Items were originally formulated in Dutch. A description of the item in English is provided.




| *Efficacy – Dealing with differences (across migration background)* | | No migr. | | | Migr. | | | Difference Δ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | λ | Rir | μ | λ | Rir | Δμ | Δλ | ΔRir |
| 1* | Adapting to other people's rules and habits | 3.04 | 0.705 | 0.49 | 3.05 | 0.672 | 0.50 | 0.01 | -0.033 | 0.01 |
| 2 | Behaving normal in an unknown environment | 3.22 | 0.675 | 0.48 | 3.27 | 0.708 | 0.53 | 0.05 | 0.033 | 0.05 |
| 3 | Adapting your language to the one who you speak with | 3.16 | 0.625 | 0.46 | 3.25 | 0.698 | 0.51 | 0.09 | 0.073 | 0.05 |
| 4 | Taking into account wishes of others when making a decision | 3.10 | 0.669 | 0.48 | 3.17 | 0.751 | 0.55 | 0.07 | 0.082 | 0.07 |

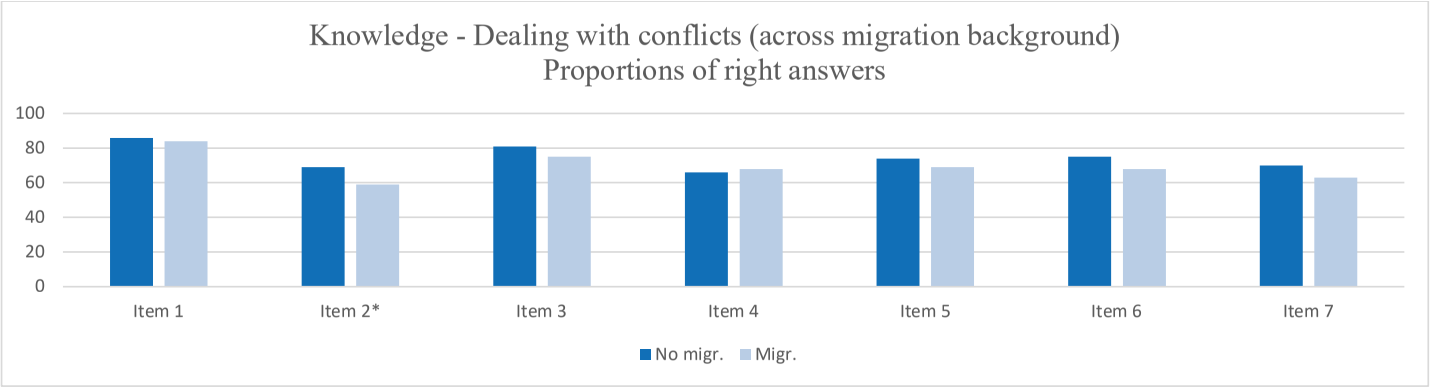
*Note.* Migr. = migration background. Items were originally formulated in Dutch. A description of the item in English is provided.






| *Knowledge – Dealing with differences (across sex)* | | Boys | | | Girls | | | Difference Δ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | λ | Rir | μ | λ | Rir | Δμ | Δλ | ΔRir |
| 1* | Recognising a prejudice in three examples | 0.62 | 0.355 | 0.20 | 0.63 | 0.505 | 0.25 | 0.01 | 0.150 | 0.05 |
| 2 | Understanding why the teacher places a classmate who behaves busy outside the classroom | 0.84 | 0.655 | 0.30 | 0.90 | 0.528 | 0.22 | 0.06 | -0.127 | -0.08 |
| 3 | Recognising discrimination in three examples | 0.88 | 0.770 | 0.34 | 0.91 | 0.747 | 0.30 | 0.03 | -0.023 | -0.04 |
| 4 | Picking the correct fact between three statements about Islam, Hinduism and Catholicism | 0.83 | 0.427 | 0.21 | 0.85 | 0.341 | 0.15 | 0.02 | -0.086 | -0.06 |
| 5* | Choosing whether a statement is discrimination, a prejudice or a fact | 0.59 | 0.487 | 0.28 | 0.63 | 0.498 | 0.25 | 0.04 | 0.011 | -0.03 |
| 6 | Deciding whether or not to take off your shoes when visiting someone where this is a habit | 0.81 | 0.555 | 0.27 | 0.87 | 0.474 | 0.20 | 0.06 | -0.081 | -0.07 |

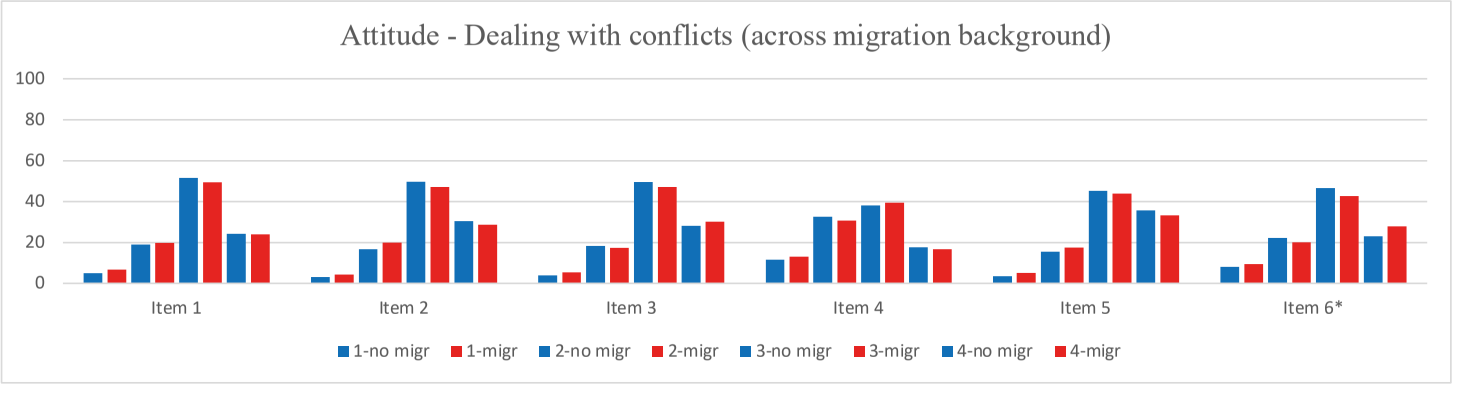
*Note.* Items were originally formulated in Dutch. A description of the item in English is provided.






| *Knowledge – Dealing with conflicts (across migration background)* | | No migr. | | | Migr. | | | Difference Δ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | λ | Rir | μ | λ | Rir | Δμ | Δλ | ΔRir |
| 1 | Making a decision with two friends when all three want something different | 0.86 | 0.651 | 0.35 | 0.84 | 0.685 | 0.37 | -0.02 | 0.034 | 0.02 |
| 2* | Deciding what you can do best when your friend is in a fight that does not seem to stop | 0.69 | 0.609 | 0.32 | 0.59 | 0.543 | 0.30 | -0.10 | -0.063 | -0.02 |
| 3 | Deciding what you can do best when you had a fight but afterwards, it appeared that you were wrong | 0.81 | 0.677 | 0.37 | 0.75 | 0.713 | 0.41 | -0.06 | 0.036 | 0.03 |
| 4 | Deciding what you can do best when your friends bully your friendly neighbour. | 0.66 | 0.418 | 0.26 | 0.68 | 0.464 | 0.29 | 0.02 | 0.046 | 0.03 |
| 5 | Deciding what you can do best when you are in a fight with a friend that only seems to become worse | 0.74 | 0.233 | 0.14 | 0.69 | 0.313 | 0.19 | -0.05 | 0.080 | 0.05 |
| 6 | Deciding what you can do best when your sister frequently skips the household chores that you need to take care of together | 0.75 | 0.627 | 0.35 | 0.68 | 0.593 | 0.35 | -0.07 | -0.034 | 0.00 |
| 7 | Deciding what you can do best when most classmates want to play a game but not everyone agrees | 0.70 | 0.633 | 0.37 | 0.63 | 0.627 | 0.38 | -0.07 | -0.006 | 0.01 |

*Note.* Migr = migration background. Items were originally formulated in Dutch. A description of the item in English is provided.




| *Attitude – Dealing with conflicts (across migration background)* | | No migr. | | | Migr. | | | Difference Δ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | λ | Rir | μ | λ | Rir | Δμ | Δλ | ΔRir |
| 1 | Taking into account the other when being into a fight | 2.95 | 0.815 | 0.62 | 2.91 | 0.763 | 0.67 | -0.04 | -0.052 | 0.05 |
| 2 | Still seeing the other in a normal way, even after disagreeing in a fight | 3.08 | 0.735 | 0.59 | 3.00 | 0.727 | 0.60 | -0.08 | -0.008 | 0.01 |
| 3 | Trying to take the other serious when being into a fight | 3.02 | 0.729 | 0.60 | 3.02 | 0.738 | 0.62 | 0.00 | 0.009 | 0.02 |
| 4 | Searching for points of agreement and disagreement when being into a fight | 2.62 | 0.703 | 0.55 | 2.60 | 0.662 | 0.59 | -0.02 | -0.041 | 0.04 |
| 5 | Getting your way when being into a fight | 3.13 | 0.757 | 0.63 | 3.05 | 0.766 | 0.63 | -0.08 | 0.009 | 0.00 |
| 6* | Willing to search a compromise when being into a fight | 2.85 | 0.592 | 0.47 | 2.89 | 0.561 | 0.50 | 0.04 | -0.031 | 0.03 |

*Note.* Migr. = migration background. Items were originally formulated in Dutch. A description of the item in English is provided.

## Appendix 2.5

*Results of Permutation Tests*

| | | Omnibus p value based on parametric chi-squared difference test | | | Omnibus p values based on nonparametric permutation method | |
|---|---|---|---|---|---|---|
| Group | Social tasks | Chisq diff | Df diff | Pr(>Chisq) | AFI. Difference | p.value |
| Sex | Attitude – acting in a socially desirable way | 200.257 | 18.000 | 0.000 | 137.802 | 0.053 |
| | Attitude – dealing with conflicts | 164.454 | 18.000 | 0.000 | 80.876 | 0.27 |
| | Attitude – dealing with differences | 1132.638 | 18.000 | 0.000 | 588.838 | 0.998 |
| | Efficacy – acting in a socially desirable way/dealing with conflicts | 413.516 | 10.000 | 0.000 | 216.566 | 1 |
| | Efficacy – dealing with differences | 33.351 | 4.000 | 0.000 | 18.636 | 0.045 |
| | Knowledge – acting democratically | 103.537 | 40.000 | 0.000 | 78.725 | 0.561 |
| | Knowledge – acting in a socially desirable way | 21.512 | 18.000 | 0.254 | 16.574 | 0.513 |
| | Knowledge – dealing with conflicts | 226.743 | 28.000 | 0.000 | 163.157 | 0.251 |
| | Knowledge – dealing with differences | 99.869 | 18.000 | 0.000 | 76.279 | 0.816 |
| SEP | Attitude – acting in a socially desirable way | 164.591 | 18.000 | 0.000 | 112.273 | 0.948 |
| | Attitude – dealing with conflicts | 114.129 | 18.000 | 0.000 | 57.182 | 0.149 |
| | Attitude – dealing with differences | 740.961 | 18.000 | 0.000 | 386.515 | 0.761 |
| | Efficacy – acting in a socially desirable way/dealing with conflicts | 263.175 | 10.000 | 0.000 | 136 | 0.529 |
| | Efficacy – dealing with differences | 19.616 | 4.000 | 0.001 | 10.999 | 0.143 |
| | Knowledge – acting democratically | 68.933 | 40.000 | 0.003 | 51.633 | 0.994 |
| | Knowledge – acting in a socially desirable way | 15.598 | 18.000 | 0.621 | 11.526 | 0.693 |
| | Knowledge – dealing with conflicts | 218.543 | 28.000 | 0.000 | 154.525 | 0.96 |
| | Knowledge – dealing with differences | 103.432 | 18.000 | 0.000 | 79.226 | 0.374 |
| Migr. | Attitude – acting in a socially desirable way | 210.179 | 18.000 | 0.000 | 141.395 | 0.012 |
| | Attitude – dealing with conflicts | 172.091 | 18.000 | 0.000 | 83.047 | 0.184 |
| | Attitude – dealing with differences | 1196.779 | 18.000 | 0.000 | 599.084 | 0.2920 |
| | Efficacy – acting in a socially desirable way/dealing with conflicts | 418.726 | 10.000 | 0.000 | 220.929 | 0.743 |
| | Efficacy – dealing with differences | 25.207 | 4.000 | 0.000 | 13.297 | 0.911 |
| | Knowledge – acting democratically | 104.332 | 40.000 | 0.000 | 78.414 | 0.598 |
| | Knowledge – acting in a socially desirable way | 19.809 | 18.000 | 0.344 | 15.149 | 0.632 |
| | Knowledge – dealing with conflicts | 218.543 | 28.000 | 0.000 | 154.525 | 0.948 |
| | Knowledge – dealing with differences | 103.432 | 18.000 | 0.000 | 79.226 | 0.38 |

**Appendix 3.1**

*Descriptive Statistics*

| | Regular primary education | Special needs primary education |
|---|---|---|
| % home language is only Dutch | 80.28% | 72.72% |
| % IQ below 75 | 0.00% | 23.18% |
| Educational advice for secondary education[9] | | |
| *Education for students with severe learning difficulty* | 0.00% | 0.63% |
| *Work/daytime activities* | 0.00% | 0.56% |
| *Practical job-oriented education* | 0.65% | 36.4% |
| *Pre-vocational secondary education (practical)* | 27.88% | 44.61% |
| *Pre-vocational secondary education (theoretical)* | 30.19% | 10.57% |
| *Senior general secondary education + pre-university secondary education* | 41.28% | 7.21% |
| Mean number of students in the entire school in 2018/19 | 258.42 | 130.83 |
| Mean urbanity (1 – 5) | 2.97 | 3.30 |
| Denomination | | |
| *Public* | 27 | 16 |
| *Catholic* | 28 | 21 |
| *Christian* | 28 | 17 |
| *Specific pedagogical/didactical focus* | 11 | 31 |

[9] In some cases, two categories were mentioned in one response. These responses were added to the lower category.

**Appendix 3.2**

*Correlation Matrix – Students in Special Needs Primary Education on the Upper Half and Students in Regular Primary Education on the Lower Half*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 0.61 | 0.20 | 0.17 | 0.10 | 0.02 | 0.11 | 0.07 | 0.06 | 0.04 | -0.01 | 0.13 | 0.04 | 0..02 | 0.00 | 0.06 | 0.07 |
| 2 | 0.45 | | 0.22 | 0.26 | 0.16 | -0.03 | 0.17 | 0.11 | 0.00 | 0.03 | 0.00 | 0.10 | 0.07 | 0.01 | 0.01 | 0.07 | 0.04 |
| 3 | 0.11 | 0.08 | | 0.58 | 0.28 | 0.04 | 0.20 | 0.18 | 0.05 | 0.04 | 0.04 | 0.05 | -0.01 | 0.11 | 0.04 | 0.04 | 0.05 |
| 4 | 0.15 | 0.16 | 0.42 | | 0.38 | -0.05 | 0.39 | 0.25 | 0.01 | 0.08 | 0.08 | 0.05 | 0.03 | 0.07 | 0.04 | 0.02 | 0.09 |
| 5 | 0.07 | 0.06 | 0.24 | 0.34 | | -0.02 | 0.45 | 0.28 | 0.01 | 0.11 | 0.15 | 0.18 | 0.10 | 0.11 | 0.14 | 0.10 | 0.10 |
| 6 | -0.05 | -0.10 | 0.01 | -0.05 | 0.02 | | -0.17 | -0.34 | 0.00 | -0.05 | -0.02 | -0.10 | 0.04 | 0.05 | 0.01 | -0.01 | -0.11 |
| 7 | 0.05 | 0.04 | 0.17 | 0.29 | 0.33 | -0.19 | | 0.47 | 0.03 | 0.10 | 0.09 | 0.15 | 0.07 | 0.08 | 0.09 | 0.06 | 0.09 |
| 8 | 0.04 | 0.00 | 0.12 | 0.18 | 0.16 | -0.32 | 0.26 | | 0.02 | 0.03 | 0.06 | 0.11 | 0.00 | 0.04 | 0.00 | -0.01 | 0.11 |
| 9 | -0.01 | 0.02 | -0.03 | 0.01 | 0.01 | -0.02 | 0.11 | -0.01 | | 0.42 | 0.46 | 0.21 | 0.09 | 0.13 | 0.06 | 0.04 | 0.16 |
| 10 | 0.02 | -0.01 | 0.07 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.40 | | 0.62 | 0.16 | 0.14 | 0.17 | 0.22 | 0.16 | 0.26 |
| 11 | -0.02 | -0.01 | 0.05 | 0.09 | 0.05 | 0.00 | 0.13 | 0.02 | 0.58 | 0.51 | | 0.13 | 0.05 | 0.16 | 0.16 | 0.13 | 0.26 |
| 12 | 0.07 | 0.04 | 0.07 | 0.08 | 0.09 | 0.01 | 0.15 | 0.01 | 0.44 | 0.52 | 0.45 | | 0.11 | 0.17 | 0.27 | 0.21 | 0.30 |
| 13 | -0.01 | 0.00 | 0.03 | 0.00 | 0.03 | -0.08 | 0.17 | 0.06 | 0.26 | 0.24 | 0.15 | 0.10 | | 0.25 | 0.32 | 0.15 | 0.00 |
| 14 | 0.05 | 0.04 | 0.02 | 0.01 | 0.02 | -0.02 | -0.01 | -0.01 | 0.05 | 0.13 | -0.02 | 0.10 | 0.24 | | 0.62 | 0.44 | 0.11 |
| 15 | 0.04 | 0.05 | -0.02 | 0.00 | -0.04 | -0.04 | 0.00 | -0.02 | 0.24 | 0.03 | 0.02 | 0.08 | 0.22 | 0.37 | | 0.46 | 0.09 |
| 16 | 0.06 | 0.04 | -0.03 | 0.02 | 0.00 | -0.06 | 0.04 | 0.02 | 0.26 | 0.16 | 0.18 | 0.15 | 0.27 | 0.46 | 0.45 | | 0.17 |
| 17 | 0.02 | 0.04 | 0.04 | 0.00 | -0.01 | -0.01 | 0.00 | 0.01 | 0.30 | 0.34 | 0.31 | 0.18 | 0.16 | 0.22 | 0.07 | 0.18 | |

*Note.* (1) Knowledge; (2) applied skills; (3) self-estimated skills; (4) attitude; (5) students' open class climate; (6) students' experienced unsafety; (7) students-teacher relationship; (8) student-student relationship; (9) teachers' attention to personal development and social competence in citizenship education; (10) teachers' attention to promoting social involvement and participation in citizenship education; (11) teachers' promotion of dealing with differences and fostering an open attitude towards differences in citizenship education; (12) teachers' perceived openness of classroom climate for discussion; (13) extent to which learning outcomes of citizenship education are monitored according to teachers; (14) extent to which a vision on citizenship education is formulated according to teachers; (15) extent to which citizenship education is translated to concrete learning goals according to teachers; (16) extent to which the curriculum of citizenship education is agreed upon throughout different educational levels according to teachers; (17) the importance of citizenship education according to teachers.

Correlations in white are not significant ($p < 0.05$); correlations in light green are significant ($p < 0.05$) and small; correlations in medium green are significant ($p < 0.05$) and medium; correlations in dark green are significant ($p < 0.05$) and large.

## Appendix 3.3

*Teacher-level Characteristics and Knowledge*

| | Knowledge – 3782 students in 179 schools | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 5<br>*B (SE)* | Model 5.1<br>*B (SE)* | Model 5.2<br>*B (SE)* | Model 5.3<br>*B (SE)* | Model 5.4<br>*B (SE)* | Model 5.5<br>*B (SE)* | Model 5.6<br>*B (SE)* | Model 5.7<br>*B (SE)* | Model 5.8<br>*B (SE)* | Model 5.9<br>*B (SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | -.04(.13)** | .05(.14) | .06(.13) | .07(.14) | -.02(.13) | -.05(.13) | -.05(.13) | -.06(.13) | -.05(.13) | .01(.14) |
| SNPE | -.02(.02) | -.23(.14) | -.23(.08)** | -.26(.11)* | -.08(.06) | -.03(.02) | -.00(.03) | .03(.05) | .03(.07) | -.12(.09) |
| *Control variables* | | | | | | | | | | |
| Girl | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* | .02(.01)* |
| Age | -.02(.01)* | -.02(.01)* | -.02(.01) | -.02(.01)* | -.02(.01)* | -.02(.01)* | -.02(.01)* | -.02(.01)* | -.02(.01)* | -.02(.01)* |
| Language spoken at home | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** | .03(.01)** |
| Advised educational level | .08(.00)** | .07(.00)** | .08(.00)** | .07(.00)** | .08(.00)** | .08(.00)** | .07(.00)** | .07(.00)** | .08(.00)** | .07(.00)** |
| IQ < 75 | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** | -.05(.02)** |
| *Teacher-level characteristics of cit. ed.* | | | | | | | | | | |
| A | .00(.02) | -.02(.02) | .00(.02) | .01(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) |
| B | .02(.01) | .01(.01) | -.02(.02) | .01(.01) | .02(.01) | .02(.01) | .01(.01) | .02(.01) | .02(.01) | .02(.01) |
| C | -.03(.02) | -.03(.02) | -.03(.02) | -.06(.02)* | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) |
| D | .04(.01)** | .05(.01)** | .05(.01)** | .05(.01)** | .03(.02) | .05(.01)** | .04(.01)** | .05(.01)** | .05(.01)** | .04(.01)** |
| E | -.00(.01) | -.00(.01) | .00(.01) | -.00(.01) | -.00(.01) | -.01(.01) | -.00(.01) | -.00(.01) | -.00(.01) | -.00(.01) |
| F | .00(.01) | -.00(.01) | .00(.01) | -.00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) |
| G | .00(.02) | .01(.02) | -.00(.02) | .00(.02) | .00(.02) | .00(.02) | .01(.02) | .02(.03) | .00(.02) | .00(.02) |
| H | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .01(.01) | .02(.01) | .02(.01) | .02(.01) |
| I | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | -.00(.02) |
| *Interaction effects* | | | | | | | | | | |
| SNPE * A | | .05(.03) | | | | | | | | |
| SNPE * B | | | .06(.02)* | | | | | | | |
| SNPE * C | | | | .06(.03)* | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SNPE * D | | | | | .02(.02) | | | | | |
| SNPE * E | | | | | | .01(.01) | | | | |
| SNPE * F | | | | | | | -.01(.01) | | | |
| SNPE * G | | | | | | | | -.04(.04) | | |
| SNPE * H | | | | | | | | | -.01(.02) | |
| SNPE * I | | | | | | | | | | .02(.02) |
| **Random part** | | | | | | | | | | |
| School-level variance | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 |
| Residual variance | .25 | .25 | .25 | .25 | .25 | .25 | .25 | .25 | .25 | .25 |
| Total variance | .32 | .32 | .32 | .32 | .32 | .32 | .32 | .32 | .32 | .32 |
| ICC | .07 | .07 | .07 | .08 | .07 | .27 | .07 | .08 | .07 | .07 |

Signif. codes: 0 '**' 0.01 '*' 0.05

*Note.* A = Attention to personal development and social competence; B = Attention to promoting social involvement and participation; C = Fostering an open attitude towards differences; D = Open classroom climate as perceived by teachers; E = Monitoring of learning outcomes of citizenship education; F = Formulation of a vision on citizenship education; G = Translation of citizenship education into learning goals; H = Agreement upon the curriculum of citizenship education; I = Importance of citizenship education according to teachers.

## Appendix 3.4

*Teacher-level Characteristics and Applied Skills*

| | Applied skills – 3782 students in 179 schools | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 5<br>*B (SE)* | Model 5.1<br>*B (SE)* | Model 5.2<br>*B (SE)* | Model 5.3<br>*B (SE)* | Model 5.4<br>*B (SE)* | Model 5.5<br>*B (SE)* | Model 5.6<br>*B (SE)* | Model 5.7<br>*B (SE)* | Model 5.8<br>*B (SE)* | Model 5.9<br>*B (SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | -.12(.17) | -.14(.18) | .00(.17) | -.04(.18) | -.08(.17) | -.13(.17) | -.14(.17) | -.13(.17) | -.12(.17 | -.11(.18) |
| SNPE | .01(.02) | .05(.18) | -.25(.10)* | -.19(.14) | -.11(.08) | .00(.02) | .04(.04) | .03(.06) | -.00(.09) | -.01(.12) |
| *Control variables* | | | | | | | | | | |
| Girl | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** |
| Age | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) |
| Language spoken at home | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** |
| Advised educational level | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** | .09(.00)** |
| IQ < 75 | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) | -.02(.02) |
| *Teacher-level characteristics of cit. ed.* | | | | | | | | | | |
| A | -.00(.02) | -.00(.03) | -.00(.02) | -.00(.02) | -.00(.02) | -.00(.02) | -.00(.02) | -.01(.02) | -.00(.02) | -.00(.02) |
| B | -.00(.02) | -.00(.02) | -.04(.02) | -.00(.02) | .00(.02) | -.00(.02) | -.00(.02) | -.00(.02) | -.00(.02) | -.00(.02) |
| C | -.01(.02) | -.01(.02) | -.01(.02) | -.03(.03) | -.01(.02) | -.01(.02) | -.01(.02) | -.01(.02) | -.01(.02) | -.01(.02) |
| D | .02(.02) | .02(.02) | .03(.02) | .03(.02) | -.00(.02) | .02(.02) | .02(.02) | .02(.02) | .02(.02) | .02(.02) |
| E | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | -.00(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| F | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | .00(.01) | -.01(.01) | -.01(.01) | -.01(.01) |
| G | .02(.03) | .02(.03) | .01(.03) | .02(.03) | .01(.03) | .02(.03) | .02(.03) | .03(.04) | .02(.03) | .02(.03) |
| H | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .01(.01) | .02(.01) | .02(.02) | .02(.01) |
| I | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.01) | .02(.02) |
| *Interaction effects* | | | | | | | | | | |
| SNPE * A | | -.01(.04) | | | | | | | | |
| SNPE * B | | | .07(.03)* | | | | | | | |
| SNPE * C | | | | .05(.03) | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SNPE * D | | | | | .05(.03) | | | | | |
| SNPE * E | | | | | | .02(.02) | | | | |
| SNPE * F | | | | | | | -.01(.01) | | | |
| SNPE * G | | | | | | | | -.02(.05) | | |
| SNPE * H | | | | | | | | | .00(.03) | |
| SNPE * I | | | | | | | | | | .00(.03) |
| **Random part** | | | | | | | | | | |
| School-level variance | .08 | .08 | .08 | .08 | .08 | .08 | .08 | .08 | .08 | .08 |
| Residual variance | .33 | .33 | .33 | .33 | .33 | .33 | .33 | .33 | .33 | .33 |
| Total variance | .41 | .41 | .41 | .41 | .41 | .41 | .41 | .41 | .41 | .41 |
| ICC | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 |

Signif. codes: 0 '**' 0.01 '*' 0.05

*Note.* A = Attention to personal development and social competence; B = Attention to promoting social involvement and participation; C = Fostering an open attitude towards differences; D = Open classroom climate as perceived by teachers; E = Monitoring of learning outcomes of citizenship education; F = Formulation of a vision on citizenship education; G = Translation of citizenship education into learning goals; H = Agreement upon the curriculum of citizenship education; I = Importance of citizenship education according to teachers.

## Appendix 3.5

*Teacher-level Characteristics and Self-estimated Skills*

| | Self-estimated skills – 3782 students in 179 schools | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 5<br>*B (SE)* | Model 5.1<br>*B (SE)* | Model 5.2<br>*B (SE)* | Model 5.3<br>*B (SE)* | Model 5.4<br>*B (SE)* | Model 5.5<br>*B (SE)* | Model 5.6<br>*B (SE)* | Model 5.7<br>*B (SE)* | Model 5.8<br>*B (SE)* | Model 5.9<br>*B (SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | .09(.14) | .21(.16) | .11(.71) | .06(.15) | .09(.15) | .10(.14) | .12(.15) | .10(.15) | .11(.15) | .12(.15) |
| SNPE | -.08(.02)** | -.40(.15)* | -.10(.09) | -.00(.12) | -.06(.07) | -.06(.02)** | -.12(.03)** | -.11(.05) | -.14(.07) | -.13(.10) |
| *Control variables* | | | | | | | | | | |
| Girl | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** | .04(.01)** |
| Age | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| Language spoken at home | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* | -.03(.01)* |
| Advised educational level | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** | .01(.00)** |
| IQ < 75 | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) |
| *Teacher-level characteristics of cit. ed.* | | | | | | | | | | |
| A | -.03(.02) | -.06(.02)* | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) |
| B | .01(.01) | .00(.01) | .00(.02) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| C | .03(.02) | .03(.02) | .03(.02) | .04(.02) | .03(.02) | .03(.02) | .03(.02) | .03(.02) | .03(.02) | .03(.02) |
| D | .03(.01) | .03(.01)* | .03(.01)* | .03(.01) | .03(.02) | .03(.01)* | .03(.01)* | .03(.01) | .03(.01) | .03(.01) |
| E | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .01(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) |
| F | .02(.01)** | .02(.01)** | .02(.01)** | .02(.01)** | .02(.01)** | .02(.01)** | .01(.01) | .02(.01)** | .02(.01)** | .02(.01)** |
| G | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.03(.02) | -.04(.02) | -.04(.03) | -.03(.02) | -.03(.02) |
| H | -.02(.01) | -.01(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.02(.01) | -.01(.01) | -.02(.01) | -.02(.01) | -.02(.01) |
| I | .01(.01) | .01(.01) | .01(.01) | .00(.01) | .01(.01) | .00(.01) | .01(.01) | .01(.01) | .00(.01) | -.00(.02) |
| *Interaction effects* | | | | | | | | | | |
| SNPE * A | | .07(.03)* | | | | | | | | |
| SNPE * B | | | .01(.02) | | | | | | | |
| SNPE * C | | | | -.02(.03) | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SNPE * D | | | | | -.01(.03) | | | | | |
| SNPE * E | | | | | | -.02(.02) | | | | |
| SNPE * F | | | | | | | .02(.01) | | | |
| SNPE * G | | | | | | | | .02(.04) | | |
| SNPE * H | | | | | | | | | .02(.02) | |
| SNPE * I | | | | | | | | | | .01(.02) |
| **Random part** | | | | | | | | | | |
| School-level variance | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 |
| Residual variance | .30 | .30 | .30 | .30 | .30 | .30 | .30 | .30 | .30 | .30 |
| Total variance | .36 | .36 | .36 | .36 | .36 | .36 | .36 | .36 | .36 | .36 |
| ICC | .04 | .04 | .04 | .04 | .04 | .04 | .03 | .04 | .04 | .04 |

Signif. codes: 0 ‘**’ 0.01 ‘*’ 0.05

*Note.* A = Attention to personal development and social competence; B = Attention to promoting social involvement and participation; C = Fostering an open attitude towards differences; D = Open classroom climate as perceived by teachers; E = Monitoring of learning outcomes of citizenship education; F = Formulation of a vision on citizenship education; G = Translation of citizenship education into learning goals; H = Agreement upon the curriculum of citizenship education; I = Importance of citizenship education according to teachers.

## Appendix 3.6

*Teacher-level Characteristics and Attitude*

| | Attitude – 3782 students in 179 schools | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 5<br>*B (SE)* | Model 5.1<br>*B (SE)* | Model 5.2<br>*B (SE)* | Model 5.3<br>*B (SE)* | Model 5.4<br>*B (SE)* | Model 5.5<br>*B (SE)* | Model 5.6<br>*B (SE)* | Model 5.7<br>*B (SE)* | Model 5.8<br>*B (SE)* | Model 5.9<br>*B (SE)* |
| **Fixed part** | | | | | | | | | | |
| (Intercept) | .04(.12) | .08(.13) | .11(.12) | .01(.13) | .04(.12) | .03(.12) | .04(.12) | .05(.12) | .03(.12) | .15(.13) |
| SNPE | -.05(.01)** | -.15(.13) | -.21(.08)* | .02(.10) | -.05(.06) | -.05(.02)** | -.06(.03)* | -.08(.05) | -.02(.06) | -.28(.08)** |
| *Control variables* | | | | | | | | | | |
| Girl | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** | .08(.01)** |
| Age | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| Language spoken at home | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) |
| Advised educational level | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* | .01(.00)* |
| IQ < 75 | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) | .00(.02) |
| *Teacher-level characteristics of cit. ed.* | | | | | | | | | | |
| A | -.03(.02)* | -.04(.02)* | -.03(.02)* | -.03(.02)* | -0.03(.02)* | -.03(.02)* | -.03(.02)* | -.03(.02) | -.03(.02)* | -.03(.02) |
| B | -.00(.01) | -.00(.01) | -.03(.02) | -.00(.01) | -.00(.01) | -.00(.01) | -.00(.01) | -.00(.01) | -.00(.01) | -.00(.01) |
| C | .04(.02)* | .04(.02)* | .04(.02)* | .05(.02)* | .04(.02)* | .04(.02)* | .04(.02)* | .04(.02)* | .04(.02)* | .04(.02)* |
| D | .03(.01)* | .03(.01)* | .03(.01)* | .02(.01)* | .03(.02) | .03(.01)* | .03(.01)* | .03(.01)* | .03(.01)* | .02(.01) |
| E | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | -.00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) |
| F | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) | .01(.01) |
| G | -.00(.02) | -.00(.02) | -.01(.02) | -.00(.02) | -.00(.02) | -.00(.02) | -.01(.02) | -.01(.03) | -.00(.02) | -.01(.02) |
| H | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) | -.01(.01) |
| I | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | .00(.01) | -.03(.01) |
| *Interaction effects* | | | | | | | | | | |
| SNPE * A | | .02(.03) | | | | | | | | |
| SNPE * B | | | .04(.02)* | | | | | | | |
| SNPE * C | | | | -.02(.02) | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SNPE * D | | | | | .00(.02) | | | | | |
| SNPE * E | | | | | | .00(.01) | | | | |
| SNPE * F | | | | | | | .01(.01) | | | |
| SNPE * G | | | | | | | | .02(.03) | | |
| SNPE * H | | | | | | | | | -.01(.02) | |
| SNPE * I | | | | | | | | | | .06(.02)** |
| **Random part** | | | | | | | | | | |
| School-level variance | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 | .06 |
| Residual variance | .23 | .23 | .23 | .23 | .23 | .23 | .23 | .23 | .23 | .23 |
| Total variance | .29 | .29 | .29 | .29 | .29 | .29 | .29 | .29 | .29 | .29 |
| ICC | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .07 |

Signif. codes: 0 '**' 0.01 '*' 0.05

*Note*. A = Attention to personal development and social competence; B = Attention to promoting social involvement and participation; C = Fostering an open attitude towards differences; D = Open classroom climate as perceived by teachers; E = Monitoring of learning outcomes of citizenship education; F = Formulation of a vision on citizenship education; G = Translation of citizenship education into learning goals; H = Agreement upon the curriculum of citizenship education; I = Importance of citizenship education according to teachers.

**Appendix 4.1**

*Results of Exploratory Factor Analysis*

| Fundamental democratic value | Factor > Item ∨ | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. |
|---|---|---|---|---|---|---|---|---|---|
| Equality | | | | | | | | | |
| | 1 | | | | | | **0.577** | | |
| | 2 | | | | | | 0.380 | | **0.499** |
| | 3 | | | | | | **0.715** | | |
| | 4 | | | | | | **0.615** | | |
| Understanding of others | | | | | | | | | |
| | 1 | | | | | **0.688** | | | |
| | 2 | | | | | **0.734** | | | |
| | 3 | | | | | **0.746** | | | |
| | 4 | | | | | **0.710** | | | |
| Freedom of speech | | | | | | | | | |
| | 1 | | | | | | | **0.446** | |
| | 2 | | | | | | | **0.394** | |
| | 3 | | | | | | | **0.488** | |
| | 4 | | | | | | | **0.578** | |
| | 5 | | | | | | | **0.643** | |
| Autonomy | | | | | | | | | |
| | 1 | **0.587** | | | | | | | |
| | 2 | **0.658** | | | | | | | |
| | 3 | **0.553** | | | | | | | |
| | 4 | 0.412 | | | | | | | **0.884** |
| Tolerance | | | | | | | | | |
| | 1 | **0.599** | | | | | | | |
| | 2 | **0.555** | | | | | | | |
| | 3 | **0.479** | | | | | | | |
| | 4 | **0.575** | | | | | | | |
| Rejection of intolerance | | | | | | | | | |
| | 1 | | | **0.779** | | | | | |
| | 2 | | | **0.840** | | | | | |
| | 3 | | | **0.778** | | | | | |
| | 4 | | | **0.801** | | | | | |
| Rejection of discrimination | | | | | | | | | |
| | 1 | | **0.817** | | | | | | |
| | 2 | | **0.852** | | | | | | |
| | 3 | | **0.827** | | | | | | |
| | 4 | | **0.815** | | | | | | |
| Sense of responsibility | | | | | | | | | |
| | 1 | | | | **0.515** | | | | |
| | 2 | | | | **0.616** | | | | |

| | |
|---|---|
| 3 | **0.679** |
| 4 | **0.559** |
| 5 | **0.615** |

*Note.* The highest factor loadings are depicted in bold and, in the case the intended scale does not cluster, second-highest factor loadings are depicted.

**Appendix 4.2**

*Assessment of Measurement Invariance across Sex, Migration Background, Home Language and Educational Sector*

| | $\chi^2$ | df | P | CFI | RMSEA | CI RMSEA | Anova | $\Delta\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|---|
| Sex (N = 1647 boys versus N = 1583 girls) | | | | | | | | | |
| Configural | 3036.3 | 998 | 0.00*** | 0.975 | 0.045 | [0.043;0.046] | - | - | - |
| Threshold | 3052.0 | 1032 | 0.00*** | 0.975 | 0.044 | [0.043;0.045] | Configural | 39.446 | 0.2395 |
| Metric | 3126.9 | 1058 | 0.00*** | 0.975 | 0.043 | [0.042;0.045] | Threshold | 69.039 | <0.001 |
| Partial metric | 3075.9 | 1053 | 0.00*** | 0.975 | 0.043 | [0.042;0.045] | Threshold | 29.608 | 0.1001 |
| Migration background (N = 3047 without migration background versus N = 231 with migration background) | | | | | | | | | |
| Configural | 3248.5 | 998 | 0.00*** | 0.976 | 0.042 | [0.040;0.043] | - | - | - |
| Threshold | 3262.6 | 1032 | 0.00*** | 0.976 | 0.041 | [0.039;0.042] | Configural | 35.577 | 0.3940 |
| Metric | 3281.3 | 1058 | 0.00*** | 0.977 | 0.039 | [0.038;0.041] | Threshold | 25.111 | 0.5127 |
| Scalar | 3310.9 | 1084 | 0.00*** | 0.979 | 0.037 | [0.036;0.039] | Metric | 26.250 | 0.4494 |
| Home language (N = 2667 Dutch or dialect of Dutch versus N = 611 other home language) | | | | | | | | | |
| Configural | 3016.6 | 998 | 0.00*** | 0.975 | 0.044 | [0.042;0.045] | - | - | - |
| Threshold | 3033.8 | 1032 | 0.00*** | 0.975 | 0.043 | [0.042;0.044] | Configural | 39.378 | 0.2418 |
| Metric | 3073.6 | 1058 | 0.00*** | 0.976 | 0.042 | [0.041;0.043] | Threshold | 35.233 | 0.1067 |
| Scalar | 3219.3 | 1084 | 0.00*** | 0.976 | 0.041 | [0.040;0.043] | Metric | 76.475 | <0.001 |
| Partial scalar | 3167.8 | 1082 | 0.00*** | 0.977 | 0.041 | [0.039;0.042] | Metric | 54.653 | <0.001 |
| Educational sector (N = 2973 regular primary education versus N = 305 special needs primary education) | | | | | | | | | |
| Configural | 3240.4 | 998 | 0.00*** | 0.976 | 0.042 | [0.041;0.044] | - | - | - |
| Threshold | 3250.8 | 1032 | 0.00*** | 0.976 | 0.042 | [0.040;0.043] | Configural | 33.005 | 0.5163 |
| Metric | 3300.7 | 1058 | 0.00*** | 0.976 | 0.041 | [0.039;0.042] | Threshold | 39.802 | 0.04079 |
| Partial metric | 3275.6 | 1056 | 0.00*** | 0.976 | 0.041 | [0.039;0.042] | Threshold | 28.153 | 0.25363 |
| Grade (N = 1367 grade 5 versus N = 1606 grade 6) | | | | | | | | | |
| Configural | 2979.8 | 998 | 0.00*** | 0.972 | 0.046 | [0.044;0.047] | - | - | - |
| Threshold | 2994.4 | 1032 | 0.00*** | 0.972 | 0.045 | [0.044;0.047] | Configural | 37.252 | 0.3218 |
| Metric | 3020.1 | 1058 | 0.00*** | 0.973 | 0.044 | [0.043;0.045] | Threshold | 26.954 | 0.4118 |
| Scalar | 3105.4 | 1084 | 0.00*** | 0.973 | 0.043 | [0.042;0.044] | Metric | 68.813 | <0.001 |
| Partial scalar | 3094.6 | 1080 | 0.00*** | 0.974 | 0.043 | [0.042;0.044] | Metric | 66.607 | <0.001 |

**Appendix 4.3**

*Comparison of Latent Means across Students With and Without a Migration Background*

| Fundamental democratic values | Standardised mean difference across migration background | P-value |
|---|---|---|
| Equality | ***-0.248*** | 0.024 |
| Understanding of others | 0.018 | 0.835 |
| Freedom of speech | **-0.294** | 0.005 |
| Autonomy | -0.203 | 0.127 |
| Tolerance | ***-0.227*** | 0.048 |
| Rejection of intolerance | **-0.300** | 0.000 |
| Rejection of discrimination | ***-0.307*** | 0.019 |
| Sense of responsibility | -0.133 | 0.233 |

*Note*. **p Δ < 0.01 (bold)**; ***p Δ < 0.05 (bold and italic).*** Standardised mean difference for students with a migration background as opposed to students without a migration background (reference group). The results indicate that students with a migration background score 0.248 units lower on equality; 0.294 lower on freedom of speech, 0.227 lower on tolerance; 0.300 lower on the rejection of intolerance; and 0.307 lower on the rejection of discrimination. There is no difference between students with and without migration backgrounds in understanding others, autonomy, and sense of responsibility.

**Summary**

A resilient democracy cannot be taken for granted but must be actively maintained by the citizens who are part of that democracy. Citizens can do so by respecting democratic processes (e.g., the rule of law) and underscoring fundamental democratic values (e.g., equality and non-discrimination). Many European countries stipulate by law that schools pay attention to this via citizenship education. In this respect, citizenship education refers to the fostering of citizenship competences among students that does not (only) takes place in a separate course but (also) in the broader school culture and general approach to education. Citizenship education is part of the socialisation function of education and aims to equip young citizens with the competences needed to participate in a democratic, pluriform society. Both socialisation and qualification (in which students are prepared for further education or participation in the economy and labour market) are important functions of education.

In line with the growing acknowledged importance of citizenship education, several studies have been dedicated to understanding the characteristics of effective citizenship education. For example, the importance of an open classroom climate for discussion for acquiring citizenship competences has been repeatedly demonstrated. However, scientific knowledge of what works in citizenship education remains scarce. This contrasts with the qualification domain of education, where school effectiveness has been studied more extensively. One of the findings to emerge from that line of research is that applying an output-driven approach predominantly leads to positive effects on students' cognitive learning outcomes (e.g., in the domains of language and mathematics). In an output-driven approach, students' learning outcomes are used in a systematic way to gain insight into the extent to which learning goals have been achieved.

Despite essential features of citizenship education that make the use of an output-driven approach less self-evident, such as the normative notion of citizenship competences (indicating that there are different perceptions on what is meant by (sufficient mastery of) citizenship competences) and the limited availability of reliable and valid measurement instruments for schools, an exploration of the possibilities for output-driven citizenship education is worthwhile. Following this premise, this dissertation examined to what extent standardised measurement instruments are able to provide for the psychometric quality and data allowing for insight into civic learning outcomes, and how civic learning outcomes can be used to improve the quality of citizenship education in primary education.

The **first study** focused on an analysis of measurement invariance as one of the understudied components of the psychometric quality of a standardised measurement

instrument. An analysis of measurement invariance provides insight into the extent to which a construct is perceived similarly in different groups. In that respect, standardised measurement instruments for citizenship competences are often used to compare different groups of learners, for example, based on sex, socioeconomic position, and migration background. Such comparative research is relevant to the quality of citizenship education because it provides insight into the extent to which students, regardless of their background characteristics, have an equal opportunity of acquiring learning outcomes. However, groups of students, for example, boys and girls, may systematically perceive certain constructs or questions in the measurement instrument differently. This hinders a meaningful comparison of results between, in this case, boys and girls. An analysis of measurement invariance reveals whether and to what extent this is the case. This study examined the measurement invariance of a widely used standardised measurement instrument for citizenship competences by looking at sex, socioeconomic position, and migration background.

The **results of the first study** show that two-thirds of the measurement instrument (i.e., 23 out of 36 scales) meets the psychometric standards concerning measurement invariance. This part, therefore, allows for comparing the results between students based on sex, socioeconomic position, and migration background. For the part that was not sufficiently measurement invariant, caution is needed when comparing results across different groups of students. Exploratory analyses can highlight which questions need further development to meet the standards for measurement invariance. In the meantime, however, these questions can still be used to describe the citizenship competences of a group, grade or school as a whole. This study not only provides insight into the psychometric quality of a widely used standardised measurement instrument for citizenship competences, but also underlines the importance of periodic analysis of measurement invariance - especially when it comes to dynamic constructs that are subject to changes over time and move with developments in society.

The **second study** provides insight into the learning outcomes and characteristics of citizenship education in regular and special needs primary education. For example, previous research on learning outcomes and characteristics of citizenship education shows that an open classroom climate for discussion promotes learning outcomes in citizenship education. However, these insights are based on studies conducted in regular primary education. Insights into learning outcomes and characteristics of citizenship education based on research conducted in special needs primary education are largely missing. The exception is a descriptive study among Dutch students in special needs primary education in 2021, which found that students in special needs primary education generally obtained lower scores on

citizenship competences than students in regular primary education. This finding was the starting point of the second study, which examined whether characteristics and citizenship competences differ across regular primary education versus special needs primary education and whether the relationship between attending special needs primary education and citizenship competences changed under the influence of characteristics of citizenship education.

The **results of the second study** show that, on average, students in special needs primary education obtain lower scores on citizenship knowledge, self-estimated citizenship skills, and citizenship attitude compared to students in regular primary education. In addition, the results show that students in special needs primary education experience the classroom climate to be more open to discussion, but also more unsafety in the school environment and more negative relationships with their teachers and peers than students in regular primary education. The relationship between attending special needs primary education and citizenship knowledge, self-estimated citizenship skills, and citizenship attitudes did not significantly change under the influence of perceived unsafety in the school environment, the teacher-student relationship and the student-student relationship, or an open classroom climate for discussion. The negative effect of attending special needs primary education on applied citizenship skills did change (i.e., it became weaker) when students experienced more positive relationships with their teacher and peers and when they experienced more openness to discussion in the classroom climate. When attending special needs primary education is seen as the moderator in an alternative interpretation of the interaction effects, the relationships between two characteristics (i.e., the teacher-student relationship and an open classroom climate) and applied skills are positive when students attend special needs primary education. However, in most cases, the relationship between characteristics of citizenship education and citizenship competences of students did not change based on whether students attended special needs primary education. This study is one of the first empirical studies to provide insight into the learning outcomes and characteristics of citizenship education in special needs primary education. As such, it lays down a foundation for further research on output-driven citizenship education in primary education.

The **third study** described the development of a new measurement instrument to measure students' attitudes towards fundamental democratic values. Underscoring fundamental democratic values, such as equality and non-discrimination, is important for the perseverance of democracy and helps different (groups of) individuals to live together peacefully in that democracy. The universality of fundamental democratic values is reflected by their great overlap with values in influential international policy documents such as the

Universal Declaration of Human Rights. In addition, fundamental democratic values generally receive wide support among citizens. In many countries, including the Netherlands, schools carry the task of visibly and actively paying attention to fundamental democratic values in education and ensuring that education does not conflict with those values. A standardised measuring instrument gives schools insight into students' attitudes towards fundamental democratic values and helps to adapt education to what students need.

The **results of the third study** give insight into the psychometric quality of the Fundamental Democratic Values Questionnaire. In that respect, the results provide empirical evidence for the validity and reliability of the measurement instrument and demonstrate that the measurement instrument meets the minimum requirements in terms of measurement invariance quality standards. The results from a first measurement occasion using the questionnaire indicate that students in the upper grades of primary education (students aged approximately 10 to 12 years old) underscore fundamental democratic values to a great extent. Support was highest for 'the rejection of discrimination' and lowest for 'understanding of others'; 'the rejection of intolerance' manifested the most disagreeing (i.e., 'somewhat disagree' or 'completely disagree') responses. The analyses of measurement invariance allowed comparisons of mean scores across migration background. The results showed that students with a migration background underscored 'equality', 'freedom of speech', 'tolerance', 'the rejection of intolerance' and 'the rejection of discrimination' to a lesser extent than students without migration backgrounds. However, caution is preserved when interpreting these results, as no additional information was available on, for example, the country of origin or the role of level of educational attainment and socioeconomic position. In sum, this chapter described the construction of a valid and reliable standardised measurement instrument that provided valuable information on young students' attitude towards fundamental democratic values and demonstrated that students agreed with fundamental democratic values to a great extent.

The **fourth study** contains a theoretical reflection on how insight into students' civic learning outcomes in primary education can be used to improve the quality of citizenship education. To this end, a model for output-driven citizenship education has been proposed. The model consists of five building blocks derived from the scientific literature. The model starts with formulating a vision on citizenship education (first building block), from which concrete learning objectives can be derived (second building block). The learning objectives are then translated into a curriculum design (third building block), from which the learning outcomes are measured and monitored (fourth building block). On this basis, it is evaluated whether the

intended learning objectives have been sufficiently achieved and whether follow-up actions are desirable or necessary (fifth building block). An output-driven approach has mainly been applied and researched in the education domain that focuses on the qualification of students (via, for instance, teaching language and mathematics), but not in the education domain that focuses on the socialisation of students, which includes citizenship education.

The **results of the fourth study** highlight two characteristics of citizenship education that make applying an output-driven approach less self-evident. The first characteristic is the normative notion of citizenship competences, which refers to the fact that it is not fixed what is meant by (sufficient mastery of) citizenship competences. In the absence of clear guidelines on the intended learning outcomes, it is expected from schools to decide what civic learning outcomes are important to foster and when these learning outcomes are at the desired level. This requires time, effort, expertise and development work. In order to meet this challenge, it is suggested to focus on a set of widely shared fundamental democratic values which are essential for the perseverance of democracy. In this respect, the *regula aurea*, better known in practice as 'to treat others like you want others to treat you', also offers a useful starting point for output-driven citizenship education. A second characteristic of citizenship education that makes applying an output-driven approach less obvious is the limited availability of valid and reliable measurement instruments for citizenship competences. Although the third study in this dissertation contributes to this challenge by describing the construction of a new measurement instrument, the importance of the ongoing development of measurement instruments for citizenship education remains undiminished. Both characteristics come into play against the background of a young measurement tradition within citizenship education and challenges with the data literacy of school professionals.

In conclusion, the individual studies in this thesis generally confirm that standardised measurement instruments are able to provide for the psychometric quality and data that allow insight into the learning outcomes of citizenship education in primary education – given the operationalisation of knowledge, attitude and skill in the measurement instruments that were used (first, second and third studies) and that insight into learning outcomes provides an opportunity to improve the quality of citizenship education (fourth study). As such, this dissertation serves as a foundation for further research into the effects of an output-driven approach on civic learning outcomes in primary education.

## Samenvatting

Een veerkrachtige democratie is geen vanzelfsprekendheid, maar moet actief worden onderhouden door de burgers die onderdeel zijn van die democratie. Dat kunnen burgers doen door democratische processen (bijvoorbeeld de rechtsstaat) te respecteren en democratische basiswaarden (bijvoorbeeld gelijkwaardigheid en non-discriminatie) te onderschrijven. In veel Europese landen is wettelijk vastgelegd dat het onderwijs hier aandacht aan besteedt via burgerschapsonderwijs. Met burgerschapsonderwijs wordt gedoeld op de bevordering van burgerschap onder leerlingen die niet (alleen) in een vak plaatsvindt, maar (ook) in de bredere schoolcultuur en algemene benadering in het onderwijs. Burgerschapsonderwijs is onderdeel van de socialisatiefunctie van het onderwijs en is erop gericht om jonge burgers toe te rusten met de competenties die nodig zijn om deel te nemen aan een democratische, pluriforme samenleving. Zowel socialisatie als kwalificatie (waarin leerlingen worden voorbereid op vervolgonderwijs of op deelname aan de economie en arbeidsmarkt) zijn belangrijke functies van het onderwijs.

In lijn met het toegenomen inzicht dat bevordering van burgerschap belangrijk is, neemt ook het aantal studies naar de kenmerken van effectief burgerschapsonderwijs toe. Zo wijst onderzoek herhaaldelijk op het belang van een open klasklimaat voor discussie voor het leren van burgerschapscompetenties. Echter, wetenschappelijke kennis over wat werkt in de bevordering van burgerschap blijft schaars. Dit staat in contrast met het kwalificatiedomein van het onderwijs, waarin schooleffectiviteit uitvoerig is bestudeerd. Een van de bevindingen die uit dat onderzoek naar voren komt, is dat het toepassen van een resultaatgerichte aanpak overwegend leidt tot positieve effecten op de cognitieve leerresultaten van leerlingen (bijvoorbeeld op het domein van taal en rekenen). In een resultaatgerichte aanpak worden de leerresultaten van leerlingen op systematische wijze gebruikt om inzicht te verkrijgen in de mate waarin leerdoelen zijn bereikt.

Hoewel er belangrijke kenmerken van burgerschapsonderwijs zijn die vragen oproepen over de toepasbaarheid van een resultaatgerichte aanpak, zoals de normatieve aard van burgerschapscompetenties (die erop wijst dat er verschillende opvattingen bestaan over wat wordt verstaan onder (voldoende beheersing van) burgerschapscompetenties) en de beperkte beschikbaarheid van betrouwbare en valide meetinstrumenten voor scholen, is een verkenning van de mogelijkheden voor resultaatgericht burgerschapsonderwijs van belang. Vanuit die premisse is in deze dissertatie onderzocht in hoeverre gestandaardiseerde meetinstrumenten voorzien in de psychometrische kwaliteit en data die inzicht in burgerschapscompetenties

mogelijk maken, en hoe inzicht in burgerschapscompetenties kan bijdragen aan een verbetering van de kwaliteit van het burgerschapsonderwijs in het primair onderwijs.

De **eerste studie** is gefocust op een analyse van meetinvariantie als een van de onderbelichte onderdelen van de psychometrische kwaliteit van gestandaardiseerde meetinstrumenten. Een analyse van meetinvariantie geeft inzicht in de mate waarin een construct in verschillende groepen op gelijke wijze wordt waargenomen. Zo worden gestandaardiseerde meetinstrumenten voor burgerschapscompetenties vaak gebruikt om verschillende groepen leerlingen te vergelijken, bijvoorbeeld op basis van geslacht, sociaaleconomische positie en migratieachtergrond. Zulk vergelijkend onderzoek is relevant voor de kwaliteit van burgerschapsonderwijs, omdat het inzicht geeft in de mate waarin leerlingen, ongeacht hun achtergrondkenmerken, een gelijke kans hebben op het behalen van leerresultaten. Het kan echter voorkomen dat groepen leerlingen, bijvoorbeeld jongens en meisjes, bepaalde constructen of vragen uit het meetinstrument systematisch op andere wijze interpreteren. Dat belemmert een zuivere vergelijking van de resultaten tussen, in dit geval, jongens en meisjes. Een analyse van meetinvariantie maakt inzichtelijk of en in hoeverre dit aan de orde is. In deze studie is de meetinvariantie van een veelgebruikt gestandaardiseerd meetinstrument voor burgerschapscompetenties onderzocht door te kijken naar geslacht, sociaaleconomische positie en migratieachtergrond.

De **resultaten van de eerste studie** laten zien dat twee derde van het meetinstrument (i.e., 23 van de 36 schalen) voldoet aan de psychometrische standaarden die gaan over meetinvariantie. Dit deel is daarom toereikend voor een vergelijking van resultaten tussen leerlingen op basis van geslacht, sociaaleconomische positie en migratieachtergrond. Bij het deel dat niet voldoende meetinvariant was, is terughoudendheid geboden bij het vergelijken van de resultaten tussen leerlingen. Exploratieve analyses kunnen uitlichten welke vragen doorontwikkeling behoeven om aan de standaarden voor meetinvariantie te voldoen. In de tussentijd kunnen deze vragen wel worden gebruikt om een beeld te geven van de burgerschapscompetenties van een groep, leerjaar of school als geheel. Deze studie geeft niet alleen inzicht in de psychometrische kwaliteit van een veelgebruikt gestandaardiseerd meetinstrument voor burgerschapscompetenties, maar onderstreept ook het belang van een periodieke analyse van meetinvariantie – zeker als het gaat om dynamische constructen die onderhevig zijn aan veranderingen in de tijd en meebewegen met ontwikkelingen in de samenleving.

De **tweede studie** geeft inzicht in de burgerschapscompetenties en kenmerken van de bevordering van burgerschap in het reguliere basisonderwijs en speciaal (basis)onderwijs.

Eerder onderzoek naar burgerschapscompetenties en kenmerken van burgerschapsonderwijs laat bijvoorbeeld zien dat een open klasklimaat voor discussie burgerschapscompetenties bevordert. Echter, deze inzichten zijn gebaseerd op studies die zijn uitgevoerd in het reguliere onderwijs. Inzicht in burgerschapscompetenties en kenmerken van burgerschapsonderwijs gebaseerd op onderzoek uitgevoerd in het speciaal (basis)onderwijs ontbreekt grotendeels. De uitzondering is een beschrijvend onderzoek onder Nederlandse leerlingen in het speciaal (basis)onderwijs in 2021, waaruit bleek dat leerlingen in het speciaal (basis)onderwijs over het algemeen lagere scores behaalden op burgerschapscompetenties dan leerlingen in het reguliere basisonderwijs. Deze bevinding was het uitgangspunt van de tweede studie, waarin werd onderzocht of schoolkenmerken van burgerschapsonderwijs en burgerschapscompetenties verschillen tussen het reguliere basisonderwijs versus het speciaal (basis)onderwijs. Ook is onderzocht of schoolkenmerken van burgerschapsonderwijs (bijvoorbeeld een open klasklimaat voor discussie) invloed hebben op de relatie tussen het volgen van speciaal (basis)onderwijs en burgerschapscompetenties.

De **resultaten van de tweede studie** laten zien dat leerlingen in het speciaal (basis)onderwijs gemiddeld lagere scores behalen op burgerschapskennis, zelf ingeschatte burgerschapsvaardigheden, en burgerschapsattitude in vergelijking met leerlingen uit het reguliere basisonderwijs. Daarnaast laten de resultaten zien dat leerlingen in het speciaal (basis)onderwijs meer openheid voor discussie in het klasklimaat ervaren, maar ook meer onveiligheid in de schoolomgeving en negatievere relaties met hun leerkrachten en medeleerlingen in vergelijking met leerlingen in het reguliere basisonderwijs. De relatie tussen het volgen van speciaal (basis)onderwijs en burgerschapskennis, zelf ingeschatte burgerschapsvaardigheden en burgerschapsattitude veranderde niet significant onder invloed van ervaren onveiligheid in de schoolomgeving, de leerkracht-leerlingrelatie en de leerling-leerlingrelatie, of een open klasklimaat voor discussie. De negatieve relatie tussen het volgen van speciaal (basis)onderwijs en toegepaste burgerschapsvaardigheden werd zwakker wanneer leerlingen positieve relaties met hun leerkracht en medeleerlingen ervaarden, en wanneer ze meer openheid voor discussie in het klasklimaat ervoeren. Een alternatieve interpretatie van de interactie-effecten, waarbij het volgen van speciaal (basis)onderwijs werd gezien als moderator, laat zien dat de relaties tussen twee kenmerken (namelijk de leerkracht-leerlingrelatie en een open klasklimaat voor discussie) en toegepaste burgerschapsvaardigheden positief zijn voor leerlingen in het speciaal (basis)onderwijs. In de meeste gevallen verschilde de relatie tussen kenmerken van burgerschapsonderwijs en burgerschapscompetenties echter niet tussen leerlingen in het reguliere basisonderwijs en

speciaal (basis)onderwijs. Deze studie is een van de eerste empirische studies die inzicht geeft in de burgerschapscompetenties en kenmerken van burgerschapsonderwijs in het speciaal (basis)onderwijs en legt daarmee een basis voor verder onderzoek naar resultaatgericht burgerschapsonderwijs in het basisonderwijs.

De **derde studie** beschrijft de ontwikkeling en psychometrische eigenschappen van de Fundamental Democratic Values Questionnaire. Dit meetinstrument is bedoeld voor het meten van attitudes ten opzichte van democratische basiswaarden. Het onderschrijven van democratische basiswaarden, zoals gelijkwaardigheid en non-discriminatie, is van belang voor het voortbestaan van de democratie en draagt eraan bij dat verschillende (groepen) individuen op vreedzame wijze kunnen samenleven in die democratie. Het universele karakter van democratische basiswaarden wordt weerspiegeld in het feit dat ze zijn opgenomen in invloedrijke internationale verdragen zoals het Universeel Verdrag voor de Rechten van de Mens, en ook onder burgers op brede steun kunnen rekenen. In veel landen, waaronder Nederland, hebben scholen mede als taak om in het onderwijs zichtbaar en actief aandacht te geven aan democratische basiswaarden, en ervoor te zorgen dat het onderwijs niet strijdig is met die basiswaarden. Een gestandaardiseerd meetinstrument geeft scholen inzicht in de attitude van leerlingen ten opzichte van democratische basiswaarden en helpt om het onderwijs aan te passen op de leerbehoefte van leerlingen.

De **resultaten van de derde studie** beschrijven de psychometrische eigenschappen van de Fundamental Democratic Values Questionnaire. Zo kon de validiteit en betrouwbaarheid van het meetinstrument empirisch worden onderbouwd en ondersteunen de resultaten de minimumeisen op het gebied van meetinvariantie. De eerste resultaten die met het meetinstrument zijn verzameld wijzen erop dat leerlingen in de laatste twee groepen van het primair onderwijs (circa tien tot twaalf jaar oud) democratische basiswaarden in hoge mate ondersteunen. De steun was het hoogst voor het ‘afwijzen van discriminatie’ en het laagst voor ‘begrip voor anderen’; bij het ‘afwijzen van onverdraagzaamheid’ manifesteerden zich de meeste afkeurende (i.e., ‘helemaal oneens’) antwoorden. De analyses van meetinvariantie lieten tussen leerlingen met en zonder migratieachtergrond een vergelijking van gemiddelde scores toe. De resultaten daarvan laten zien dat leerlingen met een migratieachtergrond ‘gelijkwaardigheid’, ‘vrijheid van meningsuiting’, ‘verdraagzaamheid’, ‘het afwijzen van onverdraagzaamheid’ en ‘het afwijzen van discriminatie’ in mindere mate ondersteunden dan leerlingen zonder migratieachtergrond. Echter, zonder aanvullende informatie over bijvoorbeeld het land van herkomst of over de rol van het onderwijsniveau en de sociaaleconomische positie van deze leerlingen, is voorzichtigheid geboden bij het

interpreteren van deze resultaten. Om die reden, en vanwege het belang van aanvullend toetsend onderzoek, kunnen hier vooralsnog geen verdere conclusies aan worden verbonden.

De **vierde studie** bevat een theoretische reflectie op hoe inzicht in de burgerschapscompetenties van leerlingen in het primair onderwijs kan bijdragen aan een verbetering van de kwaliteit van het burgerschapsonderwijs. Daartoe is een model voor resultaatgericht burgerschapsonderwijs voorgesteld. Het model bestaat uit vijf bouwstenen die zijn afgeleid uit wetenschappelijke literatuur. Het model start met het formuleren van een visie op burgerschapsonderwijs (eerste bouwsteen), van waaruit concrete leerdoelen kunnen worden herleid (tweede bouwsteen). De leerdoelen worden vervolgens vertaald naar een curriculumontwerp (derde bouwsteen), waarvan naderhand de leerresultaten worden gemeten en gemonitord (vierde bouwsteen). Op basis daarvan wordt geëvalueerd of de beoogde leerdoelen voldoende zijn behaald en of vervolgacties wenselijk of noodzakelijk zijn (vijfde bouwsteen). Een resultaatgerichte aanpak is tot dusver voornamelijk toegepast en onderzocht in het onderwijsdomein dat zich richt op kwalificatie (via, bijvoorbeeld, onderwijs in taal en rekenen), maar niet in het onderwijsdomein dat zich richt op socialisatie, waartoe ook burgerschapsonderwijs behoort.

De **resultaten van de vierde studie** belichten twee kenmerken van burgerschapsonderwijs die vragen oproepen over de toepasbaarheid van een resultaatgerichte aanpak. Het eerste kenmerk is de normatieve aard van burgerschapscompetenties, die verwijst naar het feit dat niet vaststaat wat onder (voldoende beheersing van) burgerschapscompetenties wordt verstaan. Wanneer heldere richtlijnen over de beoogde leerresultaten ontbreken, wordt van scholen verwacht dat zij kunnen aangeven welke leerresultaten van belang zijn en wanneer leerresultaten op het gewenste niveau zijn. Dat vraagt meer tijd, moeite, expertise en ontwikkelwerk. Om hieraan tegemoet te komen, wordt voorgesteld te focussen op een set gemeenschappelijke en breed gedragen democratische basiswaarden die essentieel zijn voor het voortbestaan van de democratie. Daarbij biedt ook de *regula aurea*, in de praktijk beter bekend als ‘behandel anderen zoals je zelf ook wil worden behandeld’, een handzaam uitgangspunt en startpunt voor resultaatgericht burgerschapsonderwijs. Een tweede kenmerk van burgerschapsonderwijs dat vragen oproept over de toepasbaarheid van een resultaatgerichte aanpak, is de beperkte beschikbaarheid van valide en betrouwbare meetinstrumenten voor burgerschapscompetenties. Hoewel de derde studie uit deze dissertatie daaraan een bijdrage levert, blijft het belang van het ontwikkelen van meetinstrumenten voor burgerschapsonderwijs onverminderd. Beide kenmerken spelen tegen de achtergrond van een

jonge meettraditie binnen burgerschapsonderwijs en uitdagingen rond datageletterdheid van schoolprofessionals.

Concluderend kan worden gesteld dat de afzonderlijke studies in deze dissertatie bevestigen dat gestandaardiseerde meetinstrumenten kunnen voorzien in de psychometrische kwaliteit en data die inzicht in burgerschapscompetenties in het primair onderwijs mogelijk maken – gegeven de operationalisering van kennis, attitude en vaardigheid in de gehanteerde meetinstrumenten (eerste, tweede en derde studie), en dat inzicht in burgerschapscompetenties een kans biedt om de kwaliteit van het burgerschapsonderwijs te verbeteren (vierde studie). Daarmee legt deze dissertatie een basis voor verder onderzoek naar de effecten van een resultaatgerichte aanpak op de burgerschapscompetenties van leerlingen in het primair onderwijs.

**Author Contributions**

CRediT author statements (https://casrai.org/credit/)

Chapter 2

LH: conceptualisation, data curation, formal analysis, investigation, methodology, writing – original draft. BZ: conceptualisation, methodology, formal analysis, data curation, investigation. AM: conceptualisation, methodology, writing – review & editing, supervision. ABD: conceptualisation, methodology, writing – review & editing, supervision.

Chapter 3

LH: conceptualisation, data curation, formal analysis, investigation, methodology, writing – original draft. ES: conceptualisation, data curation, formal analysis, investigation, methodology, writing – review & editing, project administration. AM: conceptualisation, methodology, writing – review & editing, supervision. ABD: conceptualisation, methodology, writing – review & editing, supervision, project administration.

Chapter 4

LH: conceptualisation, data curation, formal analysis, investigation, methodology, writing – original draft. AM: conceptualisation, methodology, writing – review & editing, supervision. ABD: conceptualisation, methodology, writing – review & editing, supervision.

Chapter 5

LH: conceptualisation, formal analysis, investigation, methodology, writing – original draft. AM: conceptualisation, methodology, writing – review & editing, supervision. ABD: conceptualisation, methodology, writing – review & editing, supervision.

**List of Publications**

**Peer-reviewed Publications or those in Progress**

**Hoek, L.**, Munniksma, A., & Dijkstra, A. B. (2023). Students' attitudes towards fundamental democratic values: The construction of a measurement instrument. *Manuscript revised and resubmitted.*

**Hoek, L.**, Slijkhuis, E., Munniksma, A., & Dijkstra, A. B. (2023). Understanding effective citizenship education in special needs education: Comparing regular primary education with special needs primary education. *Accepted for publication in Citizenship Teaching & Learning.*

**Hoek, L.**, Zijlstra, B., Munniksma, A., & Dijkstra, A. B. (2023). Measuring citizenship competences: Assessment of measurement invariance. *Journal of Social Science Education, 22(2).* https://doi.org/10.11576/jsse-5837

**Hoek, L.**, Munniksma, A., & Dijkstra, A. B. (2022). Improving citizenship competences: Towards an output-driven approach in citizenship education. *Journal of Social Science Education, 21(3),* 23-43. https://doi.org/10.11576/jsse-4303

**Contributions to Publications**

Slijkhuis, E., **Hoek, L.**, De Boer, H, Hemker, B. T., Van der Veen, I., Naayer, H. M., Stronkhorst, E., & Dijkstra, A. B. (2022). Burgerschap in het speciaal (basis)onderwijs. Technisch rapport Peil.Burgerschap 2021. Universiteit van Amsterdam.

Kosar-Altinyelken, H., **Hoek, L.**, & Jiang, L. (2020). Improving the psycho-social wellbeing of international students: The relevance of mindfulness. *British Journal of Guidance & Counselling, 48(4),* 524-536. https://doi.org/10.1080/03069885.2019.1600189

Coopmans, M., **Hoek, L.**, Geboers, E., & Dijkstra, A. B. (2019). Scholenpanels Burgerschap: Burgerschap meten. Alliantie Burgerschap.

**Presentations**

**Hoek, L.** (2023). Paper presentation ‘Understanding effective citizenship education in special needs education: Comparing regulary primary education with special needs primary education’. Brussel, België: Citizenship Education Conference.

De Schaepmeester, L., Teegelbeckers, J. Y., & **Hoek, L.** (2022). Paper symposium ‘Effectief burgerschapsonderwijs: De rol van leerkrachtovertuigingen, pedagogisch-didactische benaderingen en het evalueren van leerresultaten’. Hasselt, België: Onderwijs Research Dagen.

**Hoek, L.** (2022). Paper presentation ‘Improving citizenship competences: Towards an output-driven approach in citizenship education’. Porto, Portugal: JURE.

**Hoek, L.** (2022). Paper presentation ‘Measuring citizenship competences: Assessment of measurement invariance’. Cadiz, Spain: EARLI SIG 1&4 Joint Conference.

**Hoek, L.**, Krijnen, M., & Joris, M. (2021). Paper symposium ‘Measuring democratic citizenship at school: One measure fits all?’ Online: EARLI.

Krijnen, M., **Hoek, L.**, Henrichs, L., Van Vemde, L., Daas, R., & Berends, J. (2021). Poster symposium ‘Expliciete en impliciete burgerschapsvorming – De lessen, de leerlingen, de leerkracht en het land’. Online: Onderwijs Research Dagen.

**Hoek, L.** (2020). Paper presentation ‘Gender differences in citizenship competences: The role of students’ perceived openness of classroom climate’. San Francisco, United States of America: AERA [cancelled].

**About the Author**

Lianne Hoek obtained her bachelor's degrees in Dutch Language and Culture and Pedagogical Sciences (specialisation Educational Sciences) at the University of Groningen. Subsequently, she followed the research master Child Development and Education at the University of Amsterdam. Soon after graduating, she started her PhD research at the Research Institute of Child Development and Education at the University of Amsterdam. During her PhD project, she was involved in a study initiated by the Dutch Inspectorate of Education examining citizenship education in special needs primary education. In addition, she coordinated annual measurements of citizenship competences for schools and monitored the quality of measurements. Currently, Lianne works as an assistant professor at the University of Amsterdam. She is involved in an interdisciplinary research group studying education, citizenship and democracy in a digital world.