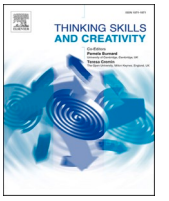
Thinking Skills and Creativity 50 (2023) 101422

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc



# Learning to think critically through Socratic dialogue: Evaluating a series of lessons designed for secondary vocational education

B.Barry Mahoney [a,*], R.Ron Oostdam [b], H.Hessel Nieuwelink [c], J.Jaap Schuitema [d]

[a] candidate at the Centre for Applied Research in Education (CARE), Amsterdam University of Applied Sciences (AUAS), Postal address: Amsterdam University of Applied Sciences, Faculty of Education, Wibautstraat 2-4, 1091 GM Amsterdam

[b] professor in Learning and Instruction and research director of the Centre for Applied Research in Education (CARE) at the Amsterdam University of Applied Sciences (AUAS). He is also professor in Learning and Instruction at the Research Institute of Child Development and Education (RICD) of the University of Amsterdam. Postal address: Amsterdam University of Applied Sciences, Faculty of Education, Wibautstraat 2-4, 1091 GM Amsterdam

[c] professor Citizenship Education at the Centre for Applied Research in Education (CARE) at the Amsterdam University of Applied Sciences (AUAS). Postal address: Amsterdam University of Applied Sciences, Faculty of Education, Wibautstraat 2-4, 1091 GM Amsterdam

[d] assistant professor at the Research Institute of Child Development and Education (RICD) of the University of Amsterdam. Postal address: University of Amsterdam, Faculty of Social and Behavioural Sciences, Nieuwe Achtergracht 166, 1018 WV, Amsterdam, the Netherlands

ARTICLE INFO

*Keywords:*
Critical thinking
Socratic dialogue
Curriculum design
Citizenship education
Secondary vocational education

ABSTRACT

This study reports the evaluation of a lesson series, developed in collaboration with teachers and experts, to generate a research-based solution for teaching critical thinking through Socratic dialogue in secondary vocational education. The lesson series has been evaluated for feasibility and tailoring to different target groups of students by five citizenship education teachers and 85 students. Data consisted of self-report questionnaires by teachers and students, complemented with observations of lessons. Results show that lessons based on Socratic dialogue for teaching critical thinking to students in secondary vocational education are considered a promising educational intervention. Teachers considered the lessons as feasible and well-tailored to their students and were able to implement the lessons in their classrooms. Students were motivated to participate, and their motivation did not decrease significantly during the lesson series. At the same time, students tended to be more motivated when the value of learning to think critically for their future profession was clearly substantiated. Practical implications from the evaluation of the lesson series were that, in addition to participating in Socratic dialogue, students need clear learning objectives and short assignments to remain active.

## 1. Introduction

According to a group of international experts in the field of education critical thinking is one of the most important learning objectives and one of the most challenging tasks for teachers (Kirschner, 2017). Aim of the present study is to generate a research-based solution for teaching critical thinking through Socratic dialogue (SD) in secondary vocational education as well as contribute to the existing body of knowledge on critical thinking and applying dialogue as a learning tool (Plomp, 2010).

Secondary vocational education in the Netherlands prepares students for higher education, future professions, and citizenship (Smulders, 2016). One of the main goals of this (democratic) citizenship education is learning to think critically. Critical thinking

* Corresponding author.
*E-mail address:* b.w.mahoney@hva.nl (B.Barry Mahoney).

https://doi.org/10.1016/j.tsc.2023.101422
Received 30 January 2023; Received in revised form 14 September 2023; Accepted 5 November 2023
Available online 7 November 2023

understood as being able to formulate a purposeful, self-regulatory judgement (Facione, 1990) means understanding the implications related to moral – and political - dilemmas in our society, which is a primary goal of citizenship education (Vaessen et al., 2021) and SD can function here as a 'communication form consistent with the pluralistic ideals of a democratic society' (Reznitskaya et al., 2009).

The effectiveness of SD as a teaching tool for critical thinking is supported by results of empirical research in higher education and secondary education (Yang, 2005; Frijters et al., 2008; McGuire, 2010; Makhene, 2019; Ho et al., 2023). To our knowledge no previous research has been done into teaching tools for critical thinking in secondary vocational education. Students in this educational training program generally have well-developed planning and training skills for a specific set of vocational skills, but at the same time are less sufficiently equipped to handle large amounts of information and high complexity levels of content knowledge (Vervoort & Elffers, 2018). Learning to think critically is therefore probably different for this group of students than for students in higher and secondary education (Mahanal et al., 2019). So, it is important to develop an educational approach that takes sufficient advantage of the fundamental features of SD while at the same time creating a learning environment appropriate to the specific characteristics and needs of these students.

To investigate the feasibility of teaching critical thinking through SD we developed a lesson series in collaboration with citizenship education teachers and experts. The final lesson series was conducted by citizenship education teachers in several classes. We report on this evaluation directed at the following three research questions: 1) Do teachers find lessons based on Socratic dialogue feasible; 2) Do teachers find lessons based on Socratic dialogue sufficiently tailored to their students, and 3) Are students motivated to participate in lessons based on Socratic dialogue? By answering these questions implications for future research into the effects of the SD on critical thinking and the educational practice are discussed.

## 2. Socratic dialogue as learning tool

Critical thinking is "a purposeful, self-regulatory judgement, which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based" (Facione, 1990, p. 6). Critical thinking is hard to learn (Van Gelder, 2005) and doesn't come naturally (Kahneman, 2012). And critical thinking is as much a skill as it is a disposition. Just learning the techniques will not be enough to foster critical thinking, developing a positive disposition, such as the willingness to apply these techniques, is just as important. Facione (1990) explores in his research a set of dispositions that are thought to be associated with developing critical thinking. Such as truth-seeking, inquisitiveness, open-mindedness, and self-confidence.

Dialogue is generally seen as a useful learning tool for critical thinking (Frijters et al., 2008). Unlike debates in which persuasion is central (Schuitema et al., 2018), the main purpose of dialogue is understanding (Daniel & Auriac, 2011; Hyde & Bineham, 2000), because "dialogic interaction is open to different perspectives. Which means that students must have the opportunity to bring in their own points of view and change the course of the discussion by bringing in their perspective" (Schuitema et al., 2018). A literature review (Abrami et al., 2015) showed significant effect sizes on learning critical thinking when students: (1) take part in dialogue, (2) in which authentic and appealing issues are discussed, and (3) a teacher plays the part of mentor, modelling dialogic skills and guiding the dialogue by providing feedback on errors and biased reasoning. Since the SD is organized through interaction, discussing authentic lived experiences, and is guided by a modelling mentor, it meets all these requirements.

Paul and Elder (2016) make a strong theoretical case for using SD for teaching critical thinking, because the questions asked by a Socratic facilitator are 'systematic, disciplined and deep' (p. 4). Also, where theory on critical thinking provides the necessary tools for knowing how we think, Socratic questions use these same tools to aim for truth and new insights and thus reaching meaningful understanding as well as better and deeper thinking.

Positive effects are reported in research into SD as a teaching tool for learning critical thinking in higher education (Yang, 2005; McGuire, 2010; Buchanan, 2012). These effects are also reported for learning critical thinking through dialogue in secondary education (Frijters et al., 2008).

### *2.1. Structure of socratic dialogue*

Socratic dialogue is: "A philosophical group dialogue in which the participants guided by a facilitator and a number of ground rules strive to reach a consensus in answering a fundamental question on the basis of one real-life example with the purpose of achieving new insights" (Knežić, 2011, p. 11). The main objective is reaching 'consensus in answering a fundamental question' (Knežić, 2011, p. 11). This means working together and co-constructing knowledge (Knežić, 2011), within a safe learning environment in which "Students can learn and flourish […] because they feel empowered to take risks by expressing their unique insights and disagreeing with others' point of view" (Gayle et al., 2013). The main aspects of SD are (1) posing a philosophical question, (2) sharing an example, (3) presenting a claim, (4) conducting a group enquiry and (5) striving for consensus.

The teacher's posed question for opening the dialogue is philosophical because of its open and fundamental character (e.g., when are you allowed to lie?). Open because the purpose of the question is aimed at clarifying the boundaries of a concept (A is lying, B is not) and fundamental because clarifying such concepts guides our way of social behaviour and moral judgement (lying is bad).

Students are asked to answer the philosophical question by sharing an example from their own life (e.g., I told my girlfriend she looked great in her new jeans, this was a lie). In this way, all students can reflect on an authentic lived experience (Altorf, 2019). After listing the examples, one example is chosen by the group for further clarification.

The student whose example has been chosen is asked to submit a claim (e.g., I was allowed to lie to my girlfriend so I wouldn't hurt her feelings), because this claim is the main object of enquiry: does this claim hold truth when submitted to rigorous questioning? After

the claim has been submitted the other students think of ways to undermine the logic of the claim by asking questions, hence starting group enquiry.

In group enquiry, the teacher asks students to try to understand each other and strive to reach consensus (Knežić, 2011), thereby motivating them to ask each other philosophical, conceptual questions (e.g., do you agree that lying to prevent hurting someone's feelings is allowed?). Although students use critical thinking skills throughout the dialogue, these are most relevant during group enquiry. Students need to submit claims, provide arguments and think of questions. Claims will be made by different students in the form of rebuttals, (counter)arguments or new perspectives. These claims should also be based on the real-life example, to which the students are continuously guided by the teacher (Knežić, 2011).

In the end the group will be bound by time (the end of the lesson), and this will force them to reach some sort of consensus. The purpose of the dialogic group enquiry is common concept formation, which is not an individual process, but directly formed by group thinking in answering questions formulated by the teacher and classmates (e.g., when trying to prevent to hurt someone's feelings, lying is sometimes allowed). This process of common concept formation can be characterized as individual thinking transforming into qualitative collective thinking (Wegerif, 2007). The group enquiry ends when the teacher summarizes what was said as common concept formation in answering the philosophical question.

### 2.2. Facilitating socratic dialogue

During group enquiry the teacher follows six intervention rules, formulated by Gustav Heckmann (1981), to ensure the dialogue to be philosophical, and aimed at reasoning. Each intervention the teacher does is (a) impartial to the contributions and characters of the students, in order to maximize their freedom to think for themselves, (b) stimulates grounding statements in the concrete real-life example, (c) promotes a willingness to reach mutual understanding, (d) focuses the groups attention back to the (fundamental) question at hand when necessary, (e) encourages the group to strive for consensus through inter-subjectively assessing all the information available and (f) directed to ensure that the dialogue stays fruitful in answering the fundamental question (Knežić, 2011, p. 23).

When using SD as a learning tool the teacher is mentor and facilitator (Abrami et al., 2015) but also acts as an example of good reasoning (Daniel & Auriac, 2011; Kirschner et al., 2018) to promote good reasoning and encourage students to use their contributions to "improve the quality of argumentation" (Wilkinson et al., 2017). Teachers need to be aware of their own level of critical thinking and their limitations. The teacher must accept and embody that "truth […] is born between people collectively searching for truth, in the process of their dialogic interaction" (Bakhtin, 1984, p. 110). In addition, teachers must familiarize themselves with the structure of the dialogue and practice themselves in asking philosophical questions.

## 3. Method

We conducted applied practice-based research (Furlong & Oancea, 2005), aimed at changing existing teaching practice to develop critical thinking skills and dispositions. For this purpose, we developed a lesson series through an iterative design process, with cooperation of educators and students as users of the lesson series. To answer our research questions, we collected data through self-report questionnaires for teachers and students, complemented with observations of lessons by the main researcher.

### 3.1. Participants

Secondary vocational education in the Netherlands consists of programs at four qualification levels (corresponding to levels 1–4 of the European Qualification Framework, European Commission, 2008). We conducted a field trial of the lesson series in the highest level (De Bruijn, 2012). Five teachers at different schools participated, teaching 85 students (see Table 1). Sampling was "strategic and purposive" (Miles et al., 2020, p. 28): citizenship teachers were selected who were available, interested in SD and critical thinking and willing to participate because they were part of the main researchers' network. Three out of five participating teachers also contributed to the development of the lesson series. Every teacher participated with one citizenship education class, in varying training programs, from beauty and styling to health care and IT. Class size was between 13 and 21 students, with disproportionate amounts of either boys or girls, depending on the type of training program (e.g., IT training programs generally have more than 90 % boys as students, the opposite being true for working in pharmacy or beauty).

Teachers were three men and two women ranging in age from mid-twenties to mid-fifties. Overall teaching experience was five years or more. Beforehand, all teachers participated in a one-day training program where they learned the basic methodology of SD

**Table 1**
Participating teachers and students, divided by training program.

| Teacher | Gender | Training program | Total students | Average age and range | Male | Female | Non-binary |
|---|---|---|---|---|---|---|---|
| T1 | male | Information technology | 21 | 18 (16–29) | 19 | 1 | 1 |
| T2 | female | Pharmacist assistant | 15 | 23 (17–37) | 2 | 13 | 0 |
| T3 | male | Beauty therapist | 13 | 18 (17–19) | 0 | 13 | 0 |
| T4 | female | Account manager | 20 | 17 (17–19) | 19 | 1 | 0 |
| T5 | male | General practitioner assistant | 16 | 19 (17–30) | 4 | 12 | 0 |
| Total | | | 85 | 19 (16–37) | 44 | 40 | 1 |

and practiced different tasks (e.g., collecting and writing down examples, provoking claims and asking Socratic questions).

This research was approved by the ethics committee of the Amsterdam University of Applied Sciences (Ref: 2020 – 111,638). All participants have given active informed consent and data files were completely anonymized.

### 3.2. Intervention

In cooperation with eight teachers who were familiar with SD and who teach citizenship education in the highest level of Dutch vocational education, a lesson series has been developed. An expert in the field of SD has been added to this group to monitor the theoretical framework of the dialogue. This group went through a construction phase (McKenney & Reeves, 2014), to develop, pilot and improve successive prototypes of the lessons (Plomp, 2010). During this process, we organized group interviews with three SD experts, to check whether the lessons met the fundamental characteristics of the format despite the pedagogical choices and practical adjustments we made.

The series of lessons is made up out of eight lessons, because this tends to be the number of lessons during one period in most schools. According to the cooperating teachers this amount provides enough opportunity to practice without becoming too intensive for secondary vocational students. To properly prepare students for SD regarding a citizenship topic (see table 3), it appeared necessary to activate their (prior) knowledge by showing a video, sharing anecdotes, or asking them to read a news article. Every lesson therefore starts with an explicit moment of activating prior knowledge of the citizenship topic, in lesson one complemented with instruction on the importance and features of critical thinking. Students also needed the satisfaction of ending the lesson with discussing answers to some of the philosophical questions raised. When evaluating the lesson, teachers are therefore motivated to ask questions about the subject matter and different perspectives on it, besides asking more general questions about the learning process. Practical implementation of the lessons ultimately showed that reaching formal consensus was not feasible within one lesson. This objective of the dialogue has therefore been adjusted: students are encouraged to strive for consensus during group enquiry, but they don't have to end with consensus in a formal conclusion.

Based on experiences during piloting the lessons, two other general measures were implemented to keep students motivated and interested. First, students were given the opportunity to write down their truth-claims and arguments before proceeding with the dialogue, giving them a sense of ownership, and shortening the listening time. And students were given the opportunity to ask procedural questions at any time during the conversation about the stage the group is currently in Knežić (2011).

To participate, students must at least gain declarative knowledge about the structure and conversational rules of SD, and they must know what a claim is and what it means to ask philosophical questions (i.e., necessary knowledge to perform the dialogue). In the lesson series, we therefore focused on three part-tasks for which students received instruction and opportunities for practicing: (a) claiming, (b) arguing and (c) questioning (see table 2). Within these three part-tasks six critical thinking skills are practiced within the context of the SD: (a) analysing, (b) interpreting, (c) inferring, (d) evaluating, (e) explaining and (f) self-regulation (Facione, 1990). These critical thinking skills are generic, because for example 'stating a claim' (e.g.: I'm hungry) is not new for students but doing so in philosophical dialogue with peers is. Critical thinking dispositions are not practiced or targeted specifically, because dispositions are trained implicitly. Through evaluation moments after SD training and with the help of modelling critical thinking dispositions by the teacher, such as 'trust in the process of reasoned inquiry' and 'open-mindedness regarding divergent world views' (Facione, 1990).

The lesson series (see Table 3) was conducted from February until mid-May 2022 in weekly citizenship education classes of one hour. Due to mandatory extracurricular activities and vacations, not all classes were conducted in a straight eight-week set. As a result, there was an extension to 10 weeks in some schools. Every lesson had one central topic (e.g., is a lifelong prison sentence justifiable?), which was chosen because of the relationship with major citizenship themes, the alignment with students' daily lives and the opportunity to formulate a central dilemma within the topic.

Every lesson was introduced to activate prior knowledge. The underlying dilemma was examined in SD by looking into real-life examples (e.g., when is punishing proportionate?) and evaluated both on process (e.g., what statements were clear, and which

**Table 2**
Overview of trained skills in part tasks (letters in bold next to the skills refer to the classification of Facione).

| Part task | Critical thinking skills | Example |
|---|---|---|
| **Claiming** | **inferring (c)** truth-claims from provided information<br>**interpreting (b)** truth claims<br>**evaluating (d)** whether claims hold truth | I was allowed to lie here.<br>Why is this called lying?<br>Was lying allowed here? |
| **Arguing** | **explaining (e)** truth-claims by providing substantiating arguments<br>**analysing (a)** arguments, their relationship with each other and with truth-claims<br>**evaluating (d)** whether provided arguments substantiate given truth-claims | I was allowed to lie here, because in that way I could prevent myself from hurting her.<br>Is not hurting the only reason to lie?<br>Did he prevent himself from hurting her feelings by lying? |
| **Questioning** | **interpreting (b)** truth-claims by asking Socratic questions<br>**explaining (e)** a claim by answering Socratic questions in a fair and accurate manner<br>**analysing (a)** and **self-regulating (f)** their own thinking when answering Socratic questions | Why do you call that remark a lie?<br>This was a lie because I didn't really mean it.<br>I believe that stating something you don't mean is a lie. |

**Table 3**
Central themes, content, and type of exercises of the eight lessons in relation to the fundamental philosophical questions.

| Lesson | Central theme of the lesson | Instruction and type of exercises | Fundamental philosophical question |
|---|---|---|---|
| (1) Instruction | Critical thinking and SD | Discussing aspects, rules and activities of SD (40 min.)<br>Short teacher guided SD (15 min.)<br>Evaluating the lesson (5 min.) | What does it mean to have freedom of choice? |
| (2) Part task practice: claiming | Voting rights for sixteen-year-olds? | Introduction central theme (5 min.)<br>Explanation claiming (5 min.)<br>Identifying claims in text (10 min.)<br>Formulating claims by students (15 min.)<br>Short teacher guided SD (15 min.)<br>Evaluating the lesson (10 min.) | Is it a good idea to grant sixteen-year-olds voting rights? |
| (3) Part task practice: arguing | Is the government entitled to interfere with our personal health and lifestyle? | Introduction central theme (5 min.)<br>Explanation arguing (5 min.)<br>Formulating arguments by students (25 min.)<br>Short teacher guided SD (15 min.)<br>Evaluating the lesson (10 min.) | What is a good reason for accepting or refusing our governments' interference with our personal health and lifestyle? |
| (4) Part task practice: questioning | To what length are we willing to treat every man as equal? | Introduction central theme (10 min.)<br>Explanation questioning (5 min.)<br>Formulating questions by students (20 min.)<br>Short teacher guided SD (15 min.)<br>Evaluating the lesson (10 min.) | What does it mean to take each other into account? |
| (5) Full dialogue | How important is neutrality regarding clothing for police officers? | Introduction central theme (5 min.)<br>Discussing rules and activities in SD (5 min.)<br>Full teacher guided SD (30 min.)<br>Evaluating the lesson (20 min.) | What does clothing tell us about who you are? |
| (6) Full dialogue | Do unhealthy choices undermine the right to medical treatment? | Introduction central theme (5 min.)<br>Discussing rules and activities in SD (5 min.)<br>Full teacher guided SD (30 min.)<br>Evaluating the lesson (20 min.) | When can you be held accountable for your own choices? |
| (7) Full dialogue | Can lifelong imprisonment be justifiable? | Introduction central theme (5 min.)<br>Discussing rules and activities in SD (5 min.)<br>Full teacher – and student guided SD (40 min.)<br>Evaluating the lesson (10 min.) | What does it mean to punish justly? |
| (8) Full dialogue | Citizenship education theme that arises from the group. | Discussing rules and activities in SD and choosing three student guides (10 min.)<br>Full student guided SD (40 min.)<br>Evaluating the lesson (10 min.) | The group chooses a fundamental question which is important to them. As a backup teachers can use the question: When are you allowed to lie? |

were not?) and content (e.g., when do you think punishment is proportionate?). The series of eight lessons starts with one instructional lesson to motivate students to engage in critical thinking through SD. The relevance of critical thinking for the personal and professional lives of students is explained explicitly because studies show (Ennis, 1989; Voogt et al., 2019) that students will better understand what critical thinking is, why it is important and how the activities they will participate in help them to think in a more critical manner when receiving explicit instruction and opportunities to practice. After this the format of the dialogue and ground rules for participating in SD (state a claim, provide arguments, listen to understand, and ask philosophical questions) are explained after which students participate in a short whole-classroom dialogue of fifteen minutes in which all aspects of SD are treated to get a first impression of the format.

Lessons two (claiming), three (arguing) and four (questioning) are aimed at practicing critical thinking skills. Every lesson starts with drawing attention to a citizenship topic (e.g.: voting rights for different age groups). Students then practice in small groups. Each lesson concludes with a short whole-classroom SD of about ten minutes to practice the trained skills in context. During this short dialogue, the citizenship topic is once again addressed and briefly examined through group enquiry, without asking students to provide real-life examples. The provided citizenship theme is the topic of enquiry. This is not comparable to participating in a whole SD yet but

aimed at practicing the trained skills in a small dialogue with peers. After the dialogue the teacher asks evaluative questions.

Lessons five through eight are complete whole-classroom dialogues, again starting with drawing attention to a citizenship domain theme and ending with evaluative questions by the teacher to question developed skills and dispositions. Lessons five and six are completely teacher-directed, but in lesson seven students become more responsible, creating space for a fully student-orientated SD in lesson eight, meaning three students will take turns as facilitator (Sfard, 1998). Main difference between these lessons and lessons 2–4, is that time needed for practicing part-tasks is now taken up by dialogue, giving students more time for group enquiry and critical thinking.

The teacher-designers unanimously believed that the final lesson series is suitable for students in vocational education and can be delivered as intended. All three experts agreed that the lesson series fulfilled the fundamental characteristics of the SD format. For an adequate and faithful implementation of the lesson series, a teacher manual has been developed with an associated training that takes a total of one day in terms of time investment (dependant on teacher's familiarity with the Socratic conversation).

### 3.3. Measurements

To ensure rich data collection we collected data through three different measurements: self-report questionnaires from the teachers, self-report questionnaires from the students, and observations of some lessons by the main researcher. These instruments were all newly developed by the main researcher.

The *self-report questionnaires for teachers* consisted of open-ended questions aimed at answering the first and second research question: 1) Do teachers find lessons based on SD feasible; 2) Do teachers find lessons based on SD sufficiently tailored to their students? The questionnaire was completed separately for each lesson. At the end of the lesson series teachers were asked about their overall impression regarding feasibility and alignment with their students. They were also asked to report practical problems they encountered (see Appendix A).

The *self-report questionnaires for students* were aimed at answering research question 3 (see Appendix C): Are students motivated to participate in lessons based on SD? The questionnaire presented claims students had to answer using a five-point Likert-scale, ranging from totally disagree (1) to totally agree (5). Students were instructed every lesson to fill in the self-report questionnaires in an online format regarding their motivation for that lesson (nine items e.g.: todays' citizenship education lesson was interesting; it was worth my time to go to citizenship education today; I did not understand what the point was of todays' citizenship education lesson). The expectation guiding these choices was that students would show decrease in motivation for the lessons, because the central part of every lesson is a guided Socratic dialogue and these kinds of dialogue can be a hard task, especially for adolescent students. Also, we expected differences in motivation for this type of task in different training programs. Some professions seem more naturally suitable for Socratic dialogue or talking amongst each other in general. Nurses for example are prescribed to talk to their patients to be able to provide the best care, but IT-specialists could in some branches of employment spend a whole day without talking to each other.

**Table 4**
Main coding labels self-report questionnaires teachers per research question.

| Research question | labels | labelled as such when…. | Examples of individual statements |
|---|---|---|---|
| Do teachers find the lessons feasible? | Execution of SD | …teachers report on how they executed SD per lesson: what went well and what was problematic. | "Guiding the dialogue was difficult, this effected my concentration." |
| | Participating in SD | …teachers report on how they witnessed students participating in SD per lesson. | "Students had the feeling they were supposed to attack or defend a statement." |
| | Evaluating the lesson | …teachers report on how they evaluated the lesson with the students at the end of every lesson. | "I ran out of time, so I chose not to do the evaluation." |
| Do teachers find the lessons sufficiently tailored to their students? | SD as educational intervention | …teachers report on the usefulness of SD for the target group and as intervention in citizenship education to foster critical thinking. | "Students said they understood the usefulness of the SD for their future profession." |
| | World of interest for students | …teachers report on how the lesson series resonated with, and were linked to the characteristics of, the target group. | "Using real-life examples really works well. Students seem to enjoy sharing their own experiences." |
| | Theme selection | …teachers report on how selected themes for lesson and SD were welcomed or rejected by students. | "This lesson was about a topic every student has heard or read about." |
| | Complexity of the SD format for students | …teachers report on the level of complexity of SD as intervention and judge if this level is suitable, too complex, or not complex enough. | "It remains really difficult. This is probably due to the lack of clear learning objectives, the slow nature of philosophic dialogue and the willingness of the students to make an effort." |
| | Development of students | …teachers report on their expected development of students due to participating in SD. | "Students learned a lot about stating truth-claims." |
| Do students find the lessons relevant and are they motivated to participate? | Teacher reports on student motivation | …teachers report on how they perceived motivation of students for SD as a method, selected themes and different interventions. | "Some students were participating enthusiastically, and some weren't. Last few weeks I've noticed it is hard to engage students to participate when they don't feel like it." |

In the questionnaire of the first lesson, using the same five-point Likert scale, three general topics were also addressed (see Appendix B): a) motivation for school in general (five items e.g.: the educational training program I attend suits me well; I enjoy going to school), b) motivation for the subject citizenship education (one item: how motivated are you for the subject citizenship education?), and c) motivation for citizenship education in general (five items e.g.: I do not understand why we have citizenship education in my educational training program; citizenship education always teaches me something new).

After the final lesson students also had to grade the whole lesson series on a 11-point numeric scale (see Appendix C), ranging from absolutely not (0) to absolutely (10), to provide information on their general motivation for the lesson series (three items e.g.: the past eight lessons citizenship education about critical thinking were worth my time).

*Observations of lessons* were carried out using a structured observation model aimed at providing context for the self-report questionnaires. Items were described in annotations (see Appendix D). Every teacher and his class were observed once, only by the main researcher, except for one because of practical complexities. As a result, a total of four lessons could be observed. Per type of lesson (i.e., instruction, part task practice, full dialogue) one class was observed. Observations were made by the principal investigator without being actively involved in any way during class.

### *3.4. Data analysis*

Teachers self-report questionnaires were coded in MAXQDA. Before analysis a Cohen's Kappa was calculated to confirm intercoder reliability. Overarching labels were assigned to teachers' individual statements by the principal investigator. This resulted in nine labels that were related to the three research questions (see table 4). Teachers' individual statements were then handed over to a second coder who had to categorize the statements using the same overarching labels. There was almost perfect intercoder agreement ($\kappa$=0.86, 95 % CI [.74, 0.98]), indicating that almost all individual teacher statements were labelled accordingly (Landis & Koch, 1977). Overall analysis of labelled statements was done vertically to establish patterns throughout the lesson series, and horizontally, to establish similarities and differences between teachers.

Data from the self-report questionnaires from students was analysed using SPSS. All eight questionnaires were combined, using unique numbers assigned to every unique respondent. Scales proved to be reliable for school motivation ($\alpha$=0.78/6 items) motivation for citizenship education ($\alpha$=0.75/5 items), motivation for separate lessons ($\alpha$=0.75/9 items) and motivation for the lesson series ($\alpha$=0.89/3 items).

General notes from the main researcher were used to accompany analysis of the self-report questionnaires from teachers and students. Annotations were checked for contradicting teachers self-report, especially regarding their own judgement of Socratic dialogic skills. Also, these annotations were used to interpret data from the self-report questionnaires from students, to check for possible explanatory factors for judgements of the lessons and motivation for participating in SD.

## 4. Results

In this result section we report our findings structured according to the main research questions (see Table 4). The first part contains qualitative results regarding RQ1: Do teachers find lessons based on Socratic dialogue feasible? The second part, also based on qualitative data, are results from RQ2: Do teachers find lessons based on Socratic dialogue sufficiently tailored to their students? The last part entails findings on student motivation based on quantitative data, regarding RQ3: Are students motivated to participate in lessons based on Socratic dialogue?.

### *4.1. Feasibility of the lessons (RQ1)*

*Execution of SD.* The structured manner of the dialogue was helpful and a challenge for teachers while executing SD$\alpha$. The rigour of the method puts a great strain on the cognitive workload of the teachers, and they reported the risk of losing focus at times when it is essential to act as a moderator. That's why teachers T2, T3 and T5 mentioned being well prepared and rested is crucial for successful execution. Still challenging was to stay in the philosophical disposition and not falling back in a classic teacher role. Three teachers (T1, T2 & T5) noticed themselves falling into a more explanatory teacher role, instead of solely guiding the dialogue, like T1 says: "I found it difficult to step out of my teacher role and into the role of facilitator." This role confusion also became clear through observations. In the first lesson all observed teachers were doubtful about their own skills, in the middle (lessons 2–4) teachers became more negative and towards the end (lessons 5–8) more positive, because they perceived their own attitude during dialogue as more facilitating. After one day of training and teaching the eight lessons all five teachers still felt they needed additional training to master execution of the dialogue as intended.

*Participating in SD.* All teachers mentioned that participation of students became richer after practicing part-tasks and having been a participant in short dialogues, as T1 says: "Lessons of the last few weeks finally seem to come together. Students have learned certain skills and now they can bring these into practice."

The main challenge for students was to refrain from debating or arguing. When participating some students felt they were being attacked when a classmate asked questions, according to teachers T2 and T4. As students took the role of facilitator classmates seemed to show two distinct behaviors. On the one hand teachers T1 and T2 reported that ownership of the dialogue was in the hands of the students, which seemed to lead to more active participation and enthusiasm. On the other hand, teacher T2 noticed students not taking their classmates leading to more hindering behaviour and the need for the teacher to step in.

Teachers T1 and T3 mention that the different phases in the full dialogue lessons didn't work as intended. It was supposed to lead

students from an abstract citizenship education theme through a SD about the underlying dilemma, back to the theme to alter or strengthen their judgement based on new thoughts.

*Evaluating the lesson.* Teachers T1 and T4 reported that evaluating at the end of the lesson and starting every lesson with looking back to the previous lesson is valuable. At the same time evaluation was skipped by teachers T1, T3 and T5 because they ran out of time. Teachers T1 and T5 didn't think the evaluating questions were suitable at the end of an intensive dialogue. Teacher T5 even asked for a different type of questioning because the chosen questions didn't challenge the students enough to evaluate their development.

*Additions from observations.* Observations also showed that part-task practices acted as a condition for more engaged participation, a clue that critical thinking dispositions of students might be shifting. After the part-task lessons students seemed to do better in recognizing claims, arguments, and Socratic questions during full dialogue, motivating them to be engaged in these part-tasks during full dialogue and as a result enriching the dialogue for the whole group. Observations also supported what teachers mentioned regarding participating in SD, more specific engaging in full dialogue after being introduced to an abstract citizenship dilemma: students seemed to get lost in the going back and forth between abstract theme and underlying dilemma (e.g.: abstract is 'police uniforms and religious symbols' versus dilemma 'what does clothing say about you?').

### *4.2. Tailoring of the lessons (RQ2)*

*SD as educational intervention.* Four teachers (T1, T2, T3 & T5) judge the dialogue as educational intervention suitable for discussing issues of citizenship education and learning critical thinking. Only teacher T4 remained sceptical, because of the nature of SD: "I don't think students in secondary vocational education have what it takes to concentrate this long during dialogue."

Teacher T3 mentioned that the highly structured SD format entailed the risk of the pace slowing down which leads to students becoming passive. Teacher T5 on the other hand reported that his students seemed supported by the highly structured form of the dialogue. The four teachers who were positive about the appropriateness of SD for their students became even more positive towards the end of the lesson series (lessons 5–8).

*World of interest of students.* Regarding the philosophical questions and real-life examples central too SD during the lessons one teacher was sceptical, two saw varying attitudes and, two were positive. According to teacher T4 the SD was too slow and boring for her students and her students didn't understand why they needed to be able to participate in SD as part of their job training. Teachers T1 and T3 noticed two different attitudes in their group: one half was interested, seemed to enjoy the activity and valued learning to think critically; the other half did not want to cooperate and didn't think the dialogue holds value for their own development as professionals in IT or beauty.

Teachers T2 and T5 reported highly interested students who enjoy participating in SD. Even when facilitation of the dialogue was less structured these students stayed interested and willing to participate. All the way through the end of the series teachers heard back that these students understood why they practiced participating in SD for learning critical thinking as part of their development as a professional in health care. Toward the end of the series students seemed even more active, according to the teachers because of the real-life examples. Students questioned each other on highly relatable topics, such as when lying is allowed.

All teachers recognize the importance of stating clear learning objectives for secondary vocational students. Students are more willing to participate in lessons when it is clear to them why they must attend the lesson as part of their job training or to develop future citizenship participation. Otherwise, as T1 says: "Students were restless and asked me what the necessity was of learning how to ask philosophical questions." There must be a clear 'for what' for every lesson, in the sense of homework, an assignment or clearer learning objectives. Students who don't understand the purpose of the dialogue are not interested enough intrinsically to participate.

Just listening and thinking could only be asked of students for a limited period. They had to become active to display goal-orientated behaviors that motivated them. Alignment increased by using educational interventions such as watching short videos and reading texts, although teacher T4 mentioned concentration issues, and teacher T2 noticed hindering low language proficiency. Teacher T3 stated that asking students to do something (e.g.: underlining claims in a text) motivates them.

*Theme selection.* Themes chosen (see table 3) for the dialogues were considered suitable by the teachers, regardless of the educational training program. Although there were differences between training programs, there was not one theme that was obviously (un) suitable (e.g.: GP-assistants enjoyed dialogue on voting rights, account managers did not). All teachers agreed that focusing dialogue on themes which were less abstract and closer to the life of their students (e.g.: 'when are you allowed to lie?') works the best. Aligning the theme of enquiry with their own lives seemed highly effective in raising the suitability of SD for secondary vocational students.

*Complexity of the SD format for students.* None of the teachers thought the dialogue was too complex for their students. But at the same time some aspects were challenging. For example, teacher T3 reported his students found the dialogue difficult because they tended to stay friendly towards each other. They did not want to be too critical and jeopardize their relationship with classmates. And teacher T4 noticed that thinking about abstract themes was challenging for her students. Some failed to grasp the complexity of themes and remained very level-headed and unwilling to perform more complex reasoning tasks. Also, T1 found it a difficult task for his group to choose interesting practical examples for group enquiry. He noticed peer pressure and the tendency of the group to choose the most humorous of examples. On the other hand, Teachers T3 and T5 reported high levels of self-regulation. Students said they enjoyed thinking and liked noticing that their opinions weren't as solid as they thought. They were able to think about their own way of thinking.

*Development of students.* Teachers all reported learning progress of their students. Teachers T2, T3 and T4 saw development in claiming, arguing, and questioning. Teacher T2 also noticed that her students could participate more actively in the dialogue. Teachers T1, T3 and T5 reported learning effects on dispositions of their students. Their students were more willing to engage in dialogue, to postpone their judgement and to ask and answer critical questions. At the same time four teachers (all but T5) noticed a distinct

relationship between the positive disposition towards Socratic dialogue and development of skills. Students who hardly participate don't seem to learn a lot, like T1 says: "When you are asked to participate in something that doesn't interest you, motivation will be very low. When a student does not want to play the game, he will not gain a lot from these lessons.". Students themselves were positive, although a few students felt that they had learned nothing and did not find the lessons useful for their future profession. Students from teachers T2, T3 and T5 were especially positive regarding their own development, they thought they learned a lot about listening, sharing one's opinion and being able to ask critical questions.

*Additions from observations.* Observations also showed some difference in engagement in the three different types of lessons (introduction, part-task practice, and full dialogue). When the dialogue was centred around real-life examples (e.g.: when are you allowed to lie; what is helping; when is a comment inappropriate?), such as in the full dialogue lessons, students became more engaged: providing more real-life examples and needing less nudging by the teacher to come up with claims or rebuttals.

### 4.3. Student motivation (RQ3)

When conducting the lesson series, there was the possibility that students would eventually find conducting a Socratic conversation tedious and, as a result, would become less motivated for the lessons. In secondary vocational education, the different target groups of students can also differ considerably in terms of their field of study and personal interests. Nurses for example are prescribed to talk to their patients to be able to provide the best care, but IT-specialists can in some branches of employment spend a whole day without talking to each other.

As the data show (table 5) students were on average motivated for their educational training program, the subject citizenship, and citizenship education in general (see table 5). All student scores for these measures are above the scale average (> 3) and show no differences between the different educational training programs. Although teachers T1 and T3 stated that half of their students were less motivated, and teacher T4 remained very sceptical, students themselves report they are motivated to take part in the lesson series (overall mean of 6.8). This seems to apply to a greater extent to GP assistants ($M = 7.92$) from teacher T5 and beauty therapists ($M = 8.00$) from teacher T3. Motivation of students for the individual lessons is on average above the scale mean and between the different courses there are hardly any differences in motivation overall. A closer inspection of the means within educational training programs reveals that pharmacist assistants and GP assistants are somewhat extra motivated for certain lessons, but these are all minor differences. Pharmacist assistants find the part tasks lessons above average motivating and GP assistants the lessons 6 to 8 with a full dialogue. Throughout the lesson series no decline of student motivation can be observed, nor difference in motivation between lessons. Only for lessons 4 and 8 are there small negative average scores for motivation for account managers and IT students respectively. This indicates that motivation for these types of lessons does not decrease for the different training programs.

## 5. Discussion

Although critical thinking is one of the most important learning objectives of contemporary education (Kirschner, 2017), it is also one of the most challenging tasks for teachers. Aim of the present study was to generate a research-based solution for teaching critical thinking through SD in secondary vocational education. A lesson series was evaluated using the following research questions: 1) Do teachers find lessons based on SD feasible; 2) Do teachers find lessons based on SD sufficiently tailored to their students, and 3) Are students motivated to participate in lessons based on SD?

**Table 5**
Results self-report questionnaires students, divided to different educational training programs (means and standard deviations between brackets); all measures with a five-point Likert scale (1 = totally disagree, 2 = disagree, 3 = don't disagree / don't agree, 4 = agree, 5 = totally agree) except motivation Socratic lessons with an eleven-point numeric scale (0 = absolutely not, 10 = absolutely).

| Measurement | N items | Information technology ($N = 21$) | Pharmacist assistant ($N = 15$) | Beauty therapist ($N = 13$) | Account manager ($N = 20$) | GP assistant ($N = 16$) | Total group ($N = 85$) |
|---|---|---|---|---|---|---|---|
| School motivation | 5 | 3.57 (0.67) | 3.82 (0.47)* | 3.10 (0.43) | 3.14 (0.52) | 3.20 (0.59) | 3.37 (0.61) |
| Motivation subject citizenship | 1 | 3.43 (1.17) | 3.73 (0.80) | 3.46 (0.67) | 3.40 (0.68) | 3.38 (0.81) | 3.47 (0.85) |
| Motivation citizenship general | 5 | 3.30 (0.96) | 3.67 (0.63) | 3.25 (0.54) | 3.36 (0.52) | 3.61 (0.64) | 3.43 (0.70) |
| Motivation Socratic lessons | 3 | 6.20 (2.42) | 6.50 (0.24) | 8.00 (1.59) * | 6.28 (2.42) | 7.92 (1.40) | 6.80 (2.17) |
| Lesson 1 Instruction | 9 | 3.06 (0.76) | 3.78 (0.40)** | 3.41 (0.41) | 3.25 (0.50) | 3.57 (0.39) | 3.38 (0.59) |
| Lesson 2 part-task claiming | 9 | 3.21 (0.51) | 3.70 (0.44)** | 3.11 (0.52) | 3.04 (0.63) | 3.26 (0.46) | 3.25 (0.56) |
| Lesson 3 part-task arguing | 9 | 3.21 (0.75) | 3.51 (0.64) | 3.21 (0.59) | 3.17 (0.34) | – | 3.27 (0.59) |
| Lesson 4 part-task questioning | 9 | 3.40 (0.79) | 3.62 (0.41)* | 3.36 (0.39) | 2.98 (0.50)* | 3.56 (0.38) | 3.35 (0.59) |
| Lesson 5 full dialogue | 9 | 3.29 (0.50) | 3.80 (0.38)** | 3.27 (0.45) | 3.19 (0.58) | 3.43 (0.41) | 3.37 (0.52) |
| Lesson 6 full dialogue | 9 | 3.25 (0.74) | 3.25 (0.92) | 3.64 (0.34)* | 3.13 (0.42) | 3.70 (0.23)** | 3.33 (0.59) |
| Lesson 7 full dialogue | 9 | 3.16 (0.88) | 3.36 (0.31) | 3.13 (0.47) | – | 3.86 (0.28)** | 3.30 (0.69) |
| Lesson 8 full dialogue | 9 | 2.73 (0.62)* | 3.33 (0.47) | 3.07 (0.54) | 3.19 (0.31) | 3.97 (0.51)** | 3.13 (0.59) |
| Total over lessons | | 3.16 (0.69) | 3.54 (0.50) | 3.28 (0.46) | 3.14 (0.47) | 3.62 (0.38)* | 3.30 (0.59) |

Note 1= Difference between mean of the group and mean of the whole group expressed as pooled variance is medium to large, 0.5 to 0.8 (*) or large >0.8 (**). Note 2= questionnaire lesson seven by account managers and lesson three by GP-assistants were not submitted, corresponding lesson was delivered.

Concerning the first research question we can conclude that teachers find the lessons feasible and demanding at the same time. The structured way of the dialogue proved useful, but only if teachers were equipped and sufficiently prepared to guide the dialogue so they could act as examples of good reasoning themselves, which is according to empirical evidence from other studies an essential feature of teaching for critical thinking (Daniel & Auriac, 2011; Kirschner et al., 2018). Teachers had to prepare themselves thoroughly for every single lesson, but for the whole lesson series teachers need to be prepared even more. Especially concerning the use of the six intervention rules, which are, also according to Knežić (2011, p. 23), a key feature of guiding SD. We found that teachers need extensive training in using the intervention rules to stay in the role of facilitator of dialogue.

Concerning the second research question we can conclude that most teachers find the lessons sufficiently tailored to their students, but lessons can be tailored some more by providing more short assignments to keep students active and by helping them to handle large amounts of content knowledge, as this especially is challenging for secondary vocational students (see also Vervoort & Elffers, 2018; Mahanal et al., 2019). As results from studies in secondary education indicate (Van Gelder, 2005; Wegerif, 2007; Heijltjes et al., 2014; Volman & Ten Dam, 2015; Voogt et al., 2019) we conclude that secondary vocational students also need ample opportunities to practice critical thinking skills through exercises after receiving informed instruction. Secondly, as previous research in higher education shows, the theme of the Socratic Dialogue matters, not every theme or context is suitable for this type of dialogue (Makhene, 2019). And although Abrami et al. (2015) concluded real-life examples should be a part of the lessons serving as authentic examples for dialogue, we found this to be a key feature when designing the lessons. Examples used only worked well for keeping the dialogue fruitful if they came directly from a students' personal experience. Finally, when SD is made into a lesson series like this, the lessons suit students in secondary vocational education, although staying in dialogue seems to meet the needs of students in health care training programs more easily than students in training programs for account manager, IT, and beauty. The nature of these professions might explain the different perceived alignment with these groups, but we cannot conclude that based on our findings.

Concerning the third research question we can conclude that most students are motivated to participate. A minority of students is not easily motivated because they don't see the value of learning to think critically for their future profession or don't understand how SD will help them become better critical thinkers. Teachers notice that the dialogue does not prevent these students from becoming passive. On the other hand, teachers say this has to do with choosing the right incentives for learning, such as rewarding study credits or a grade at the end of the course. Students tend to be more willing to participate when learning objectives are clear and credits are assigned. The lesson series must have clearer learning objectives and the obligation added for students to hand in an assignment at the end of the series or the end of every lesson, because to change one's disposition towards critical thinking will be developed more when students have a clear understanding of the purpose of this type of thinking.

### *5.1. Limitations and implications for further research*

This research concerns a relatively small-scale evaluation of the lesson series. Despite the small sample of five teachers and 85 students, the participating schools were well spread over different regions in the Netherlands and the five educational programs in the study were quite diverse in educational content and student population. From this perspective, the results of this study can be regarded as an indication of the extent to which the SD can be used as a learning tool for teaching critical thinking.

Another limitation concerns the fact that the primary measures concerned self-report questionnaires, which may give a biased picture of actual perceptions and experiences. Although the self-reports were supplemented with some lesson observations, this resulted in limited qualitative data about the course of events during the lessons. Further research using observations in addition to self-reports can provide a broader comprehensive perspective on the feasibility of these types of lessons and more information about the required pedagogical skills of teachers which are needed to ensure that the dialogue proceeds adequately with different target groups of students and which adjustments may be necessary to better tailor the learning tool to the special needs from learners.

Finally, follow-up research is needed into learning effects. Although the main aim of our lesson series is to teach students in secondary vocational education to think critically at a higher level, by participating in lessons based on the SD, still little is known about the extent to which this learning effect can be achieved.

### *5.2. To conclude*

General findings of the present study indicate that lessons based on SD for teaching critical thinking to students in secondary vocational education can be considered as promising. Teachers were able to deliver the lessons to their students, although they needed training beforehand to use the dialogue in their classrooms. Students also had a positive attitude towards citizenship education and most of them were motivated to participate in SD. There were no remarkable differences between students of the different educational training programs, showing that SD can be used as a learning tool independent of educational content.

## Funding

This work was supported by the Dutch Research Council (NWO) [grant number 023.015.029].

## CRediT authorship contribution statement

**B.Barry Mahoney:** Conceptualization, Methodology, Writing – original draft, Formal analysis, Investigation, Visualization, Project administration, Funding acquisition. **R.Ron Oostdam:** Writing – review & editing, Supervision, Funding acquisition. **H.Hessel**

**Nieuwelink:** Writing – review & editing, Supervision. **J.Jaap Schuitema:** Writing – review & editing, Supervision.

## Data availability

Data will be made available on request.

## Appendix A. The self-report questionnaire for teachers

| Question | Aimed at |
|---|---|
| ***1. Questions per lesson*** | Generic |
| Date and time? | |
| Attendance, boys, girls, and others? | Generic |
| What is your general impression regarding the practical feasibility of this lesson? | RQ 2 |
| What is your general impression regarding the presumed learning outcome of this lesson? | RQ 2 |
| What parts of the lesson are practical and useful? | RQ 2 |
| What parts of the lesson are not so practical and / or not so useful? | RQ 2 |
| What was most notable during the lesson? | RQ 2 |
| What did the students think of this lesson (ask them or make an educated guess)? | RQ 2 |
| To what extent were you successful in guiding the SD? | RQ 1 |
| Do you need additional training before you can teach this lesson to your own standards of quality? | RQ 1 |
| Did you alter the lesson in any way? If yes, why, and what was the effect of this alternation? | RQ 1 |
| Any other observations or remarks? | Generic |
| | |
| ***2. Questions concerning the whole lesson series*** | |
| Student group of the eight lessons (describe extensively)? | Generic |
| Particularities influencing the integrity of the lessons? | Generic |
| How do you judge the feasibility of the lessons? | RQ 2 |
| What learning effect do you think the lessons have? | RQ 2 |

## Appendix B. The self-report questionnaire for students: generic questions after lesson one

| Question | Aimed at |
|---|---|
| ***1. Open questions generic information*** | |
| What is your first and last name? | Generic |
| What is your date of birth? | |
| What is your gender? | |
| Whit what diploma did you graduate from secondary education? | |
| Which school do you attend? | |
| Who is your citizenship teacher? | |
| What training program do you follow? | |
| ***2. Motivation for school (5-point Likert scale)*** | |
| The training program I follow suits me well. | RQ3 |
| I enjoy going to school. | |
| I would recommend this training program to new students. | |
| The lessons I take are worth my time. | |
| I am motivated to finish this training program. | |
| I often don't feel like going to school. (recode) | |
| I get good grades at school. | |
| | |
| ***3. Motivation for citizenship education (5-point Likert scale)*** | |
| Citizenship education is an enjoyable subject. | RQ3 |
| I don't understand why we are taught citizenship education. (recode) | |
| Citizenship education is important for my profession. | |
| Citizenship education is a waste of my time. (recode) | |
| During citizenship education I always learn something new. | |
| | |
| ***4. Opinion on thinking skills and dispositions (5-point Likert scale)*** | |
| I enjoy thinking in an abstract manner. | RQ3 |
| I can express my own thoughts to others. | |
| When people have different opinions, they are best of talking about something else. (recode) | |
| When you have an opinion, you should be brave enough to defend it. | |

## Appendix C. The self-report questionnaire for students per lesson (5-point Likert scale)

| Question | Aimed at |
|---|---|
| ***1. Questions per lesson (5-point Likert scale)*** | |
| Todays' citizenship education lesson was interesting. | RQ3 |
| Todays' citizenship education lesson was instructive. | |
| I was motivated for todays' citizenship education lesson. | |
| Today was a nice citizenship education lesson. | |
| It was worth my time to go to citizenship education today. | |
| Todays' citizenship education lesson had a good fit with my training program. | |
| Todays' citizenship education lesson was important for my future. | |
| Todays' citizenship education lesson was too easy for me. (recode) | |
| I didn't understand what todays' citizenship education lesson was about. (recode) | |
| ***2. Questions concerning the whole lesson series (11-point numeric scale)*** | |
| The past eight citizenship education lessons about critical thinking were worth my time. | RQ 3 |
| The past eight citizenship education lessons about critical thinking were interesting. | |
| The past eight citizenship education lessons about critical thinking have helped me become a more critical thinker. | |

## Appendix D. Structured observation model

| Question | Example annotation | Aimed at |
|---|---|---|
| ***1. Observations concerning teacher skills*** | | |
| Is the teacher capable of modelling a critical thinker?<br>How does the teacher intervene during group enquiry to keep the dialogue philosophically fruitful?<br>In what way does the teacher use Socratic questions? | Teacher explains his way of thinking about his own prejudice: "why do I think young people can't vote in a proper way?"<br>Teacher asks whether the group agrees when student A states that she: ""expressed her Brazilian identity by wearing traditional Brazilian clothing."<br>Teacher asks: "what do you mean by your identity?" Student responds with: "you know, who I am." Teacher does not ask a follow-up question. | RQ 1 |
| ***2. Observations concerning pedagogic interventions*** | | |
| Does the start of the lesson motivate students to think about the citizenship topic?<br>Is the fundamental philosophical question clear enough to all students?<br>In what way does the teacher use digital tools in relation to the white board or flip over? | The start of the lesson is watching a news-item video. This introduces the topic and students are eager to share their opinion of the topic (neutrality of police uniform).<br>One student asks the teacher what the question means, it is not yet clear to him.<br>Teacher presents the fundamental question in PowerPoint. During dialogue he must go back to the rules for dialogue and loses the question in this way. | RQ 1 |
| ***3. Observations concerning student behaviour*** | | |
| Are students able and willing to provide real-life examples?<br>Can students relate their own real-life example to the fundamental philosophical question in a truth-claim?<br>Are students willing and able to explain their thinking to their peers?<br>What is the general ambiance in the group of students? | Four out of twelve students come up with a real-life example. After specifically being asked about their examples by the teacher the other eight students can't come up with anything.<br>Student A tries to answer the fundamental philosophical question (what does clothing tell us about who you are?) with her real-life example by stating that she: "expressed her Brazilian identity by wearing traditional Brazilian clothing."<br>Student B asks why student A thinks her clothes were 'traditional Brazilian'. Student A responds by saying: "I don't know, that's what my mom said." Not enough explanation of thought is provided here.<br>Students are mostly silent and don't seem eager to speak their mind. Two out of twelve students are mostly talking and are hardly limited by the teacher. | RQ 3 |

## References


Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A Meta-analysis. *Review of Educational Research, 85*(2), 275–314. https://doi.org/10.3102/0034654314551063

Altorf, H. M. (2019). Dialogue and discussion: Reflections on a Socratic method. *Arts & Humanities in Higher Education, 18*(I), 60–75. https://doi.org/10.1177/1474022216670607

Bakhtin, M. M. (1984). *Problem of dostoevsky's poetics, 8*. Minneapolis, MN: University of Minnesota. https://doi.org/10.5749/j.ctt22727z1

Bruijn, de, E (2012). Teaching in innovative education in the Netherlands. *Teachers and Teaching, 18*(6), 637–653. https://doi.org/10.1080/13540602.2012.746499

Daniel, M. F., & Auriac, E. (2011). Philosophy, critical thinking and philosophy for children. *Educational Philosophy and Theory, 43*(5), 415–435. https://doi.org/10.1111/j.1469-5812.2008.00483.x

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*(4), 4–10. https://doi.org/10.3102/0013189X018003004

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. research findings and recommendations.* Newark, NJ: American Philosophical Association.

Frijters, S., Dam, G., ten, & Rijlaarsdam, G. (2008). Effect of dialogic learning on value-loaded critical thinking. *Learning and Instruction, 18*, 66–82. https://doi.org/10.1016/j.learninstruc.2006.11.001

Furlong, J., & Oancea, A. (2005). *Assessing quality in applied and practice-based educational research*. Oxford: Oxford University Department of Educational Studies.

Gayle, B. M., Cortez, D., & Preiss, R. W. (2013). Safe spaces, difficult dialogues, and critical thinking. *International Journal for the Scholarship of Teaching and Learning, 7*(2), 1–8. https://doi.org/10.20429/ijsotl.2013.070205

Gelder, T., van (2005). Teaching critical thinking: some lessons from cognitive science. *College Teaching, 53*(1), 41–48. https://doi.org/10.3200/CTCH.53.1.41-48

Heijltjes, A., Gog, T., van, Leppink, J., & Paas, F (2014). Improving critical thinking: Effects of dispositions and instructions on economics students' reasoning skills. *Learning and Instruction, 29*, 31–42. https://doi.org/10.1016/j.learninstruc.2013.07.003

Ho, Y. R., Chen, B. Y., & Li, M. C. (2023). Thinking more wisely: Using the Socratic method to develop critical thinking skills amongst healthcare students. *BMC Medical Education, 23*(173), 1–16. https://doi.org/10.1186/s12909-023-04134-2

Hyde, B., & Bineham, J. L. (2000). From debate to dialogue: Toward a pedagogy of nonpolarized public discourse. *Southern Journal of Communication, 65*(2–3), 208–223. https://doi.org/10.1080/10417940009373168

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books Ltd.

Kirschner, P. A. (2017). *Het voorbereiden van leerlingen op (nog) niet bestaande banen*. Heerlen: Open Universiteit.

Kirschner, P. A., Claessens, L., & Raaijmakers, S. (2018). *Op de schouders van reuzen. Inspirerende inzichten uit de cognitieve psychologie voor leerkrachten*. Meppel: Ten Brink Uitgevers. B.V.

Knežić, D. (2011). *SD and teacher-pupil interaction*. Den Haag: Eleven International Publishing.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Mahanal, S., Zubaidah, S., Sumiati, I. D., Sari, T. M., & Ismirawati, N. (2019). RICOSRE: A learning model to develop critical thinking skills for students with different academic abilities. *International Journal of Instruction, 12*(2), 417–434. https://doi.org/10.29333/iji.2019.12227a

Makhene, A. (2019). The use of the Socratic inquiry to facilitate critical thinking in nursing education. *Health SA Gesondheid, 24*, 1–6. https://doi.org/10.4102/hsag.v24i0.1224

McGuire, L. A. (2010). *Improving student critical thinking and perceptions of critical thinking through direct instruction in rhetorical analysis*. Ann Arbor: ProQuest LLC.

McKenney, S., & Reeves, T. C. (2014). Educational design research. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th Edition, pp. 131–140). New York: Springer. https://doi.org/10.1007/978-1-4614-3185-5_11.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis. A methods sourcebook*. Thousand Oaks, CA: SAGE publications.

Paul, R., & Elder, L. (2016). *The thinker's guide to socratic questioning*. Foundation for Critical Thinking Press.

Plomp, T. (2010). Educational design research: An introduction. In T. Plomp, & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 11–50). Enschede: SLO.

Reznitskaya, A., Kuo, L. J., Clark, A. M., Miller, B., Jadallah, M., Anderson, R. C., et al. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education, 39*(1), 29–48. https://doi.org/10.1080/03057640802701952

Schuitema, J., Radstake, H., Pol, J., van de, & Veugelers, W. (2018). Guiding classroom discussions for democratic citizenship education. *Educational Studies, 44*(4), 377–407. https://doi.org/10.1080/03055698.2017.1373629

Smulders, H. (2016). Al het goede komt in drieën. [All good things happen in threefold.]. *Canon beroepsonderwijs*. Den Bosch: ECBO.

Vaessen, A., Daas, R., & Nieuwelink, H. (2021). All things considered: The views of adolescents in vocational education on competing democratic values. *Citizenship, Social and Economics Education*, 1–19. https://doi.org/10.1177/20471734211050290

Vervoort, M., & Elffers, L. (2018). Het achterblijvende studiesucces van mbo-instromers in het eerste jaar van het hbo: de rol van vaardigheden, leerstijl, studiekeuze en motivatie. *TvHO, 36*(1), 23–37.

Volman, M., & Dam, G., ten (2015). Critical thinking for educated citizenship. In M. Davies, & R. Barnett (Eds.), *The palgrave handbook of critical thinking in higher education* (pp. 593–603). New York: Palgrave Macmillan. https://doi.org/10.1057/9781137378057_35.

Voogt, J. M., Veltman, M. E., & Keulen, J., van (2019). Kritisch denken als een 21ste-eeuwse vaardigheid: Veelbelovende aanpakken voor de onderwijspraktijk. [Critical thinking as a 21st century skill: Promising approaches for educational practice.]. *Pedagogisch studiën, 95*, 329–334.

Wegerif, R. (2007). *Dialogic Education and Technology. Expanding the space of learning*. New York, NY: Springer. https://doi.org/10.1007/978-0-387-71142-3

Wilkinson, I. A. G., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., et al. (2017). Toward a more dialogic pedagogy: Changing teachers' beliefs and practices through professional development in language arts classrooms. *Language and Education, 31*(1), 65–82. https://doi.org/10.1080/09500782.2016.1230129

Yang, Y. C., Newby, T. J., & Bill, R. L. (2005). Using Socratic questioning to promote critical thinking skills through asynchronous discussion forums in distance learning environments. *The American Journal of Distance Education, 19*(3), 163–181. https://doi.org/10.1207/s15389286ajde1903_4