The Quality of Instruction in Swedish Lower Secondary Language Arts and Mathematics

Michael Tengberg, Jorryt van Bommel, Marie Nilsberth, Michael Walkert & Anna Nissen

To cite this article: Michael Tengberg, Jorryt van Bommel, Marie Nilsberth, Michael Walkert & Anna Nissen (2022) The Quality of Instruction in Swedish Lower Secondary Language Arts and Mathematics, Scandinavian Journal of Educational Research, 66:5, 760-777, DOI: 10.1080/00313831.2021.1910564

To link to this article: https://doi.org/10.1080/00313831.2021.1910564

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 14 Apr 2021.

Submit your article to this journal

Article views: 1351

View related articles

View Crossmark data

Citing articles: 4 View citing articles
The Quality of Instruction in Swedish Lower Secondary Language Arts and Mathematics

Michael Tengberg, Jorryt van Bommel, Marie Nilsberth, Michael Walkert and Anna Nissen

Department of Educational Studies, Karlstad University, Karlstad, Sweden

ABSTRACT
Instructional quality is a research topic that has received increased attention over the past decades. However, despite evidence for its importance to student learning, few studies are designed to examine patterns of prevalent instruction. The present study aimed to enhance the understanding of instructional quality in Swedish lower secondary school by examining patterns of instruction in 7th grade language arts and mathematics. 274 lessons from 73 separate classrooms were video-recorded and analyzed using the Protocol for Language Arts Teaching Observation (PLATO). Findings of the study show that lessons were largely organized through either whole-class instruction or individual seatwork. Mathematics included significantly more explicit teacher scaffolding and mathematics teachers scored higher than language arts teachers on nearly all instructional dimensions. However, in both subjects, there were substantial differences between teachers, meaning that students in different classrooms received systematically different kinds of instruction. Implications for instructional development and future research are discussed.

ARTICLE HISTORY
Received 24 June 2020
Accepted 15 March 2021

KEYWORDS
Instructional quality; language arts; mathematics; PLATO; teaching

Introduction
The quality of instruction is one of the most important factors in predicting student achievement and has been found to be more significant than teacher qualifications (including certification, experience, or education), class size, and classroom climate (Allen et al., 2011; Hattie, 2009; Slater et al., 2012). Moreover, teachers vary considerably in their contribution to student learning (Haertel, 2013; Konstantopoulos & Chung, 2011; Slater et al., 2012), which, in many countries, has been found to be at least as consequential between neighboring classrooms in the same school as it is between different schools in different parts of the country (Nye et al., 2004; Slater et al., 2012). However, in what way teachers differ, and to what extent quality practices are employed in Nordic classrooms, is less known.

International large-scale assessments, such as PISA, PIRLS, and TIMSS, provide comparative data on national and international trends of student achievement in vital areas such as reading, mathematics, and science. These studies include survey instruments that map patterns of instruction in participating countries and correlate features of instruction with achievement (Blömeke &
Similarly, the TALIS study gathers comparable survey data between countries about teachers’ opinions, beliefs, and perceptions of their instruction. While these studies are valuable with respect to their size and representativeness, they provide information on a limited subset of all the practices employed by teachers. Furthermore, evidence suggests that teacher self-report data might be less reliable for certain dimensions of instructional quality (Kunter & Baumert, 2006) and do not always align with classroom observations (Goe et al., 2008). Therefore, detailed and systematic observations of teaching practices may provide a useful supplement to large-scale teacher survey reports and student achievement data in order to better understand the nature and quality of the instruction students receive in school.

In the Nordic countries, there have been few attempts at tracing prevalent patterns of instruction by observing large-scale, naturally occurring instruction. In this article, we report a video-observation study of 7th grade Swedish language arts and mathematics instruction. The aim of the study is to provide empirically-based knowledge of prevalent patterns of instruction in language arts and mathematics. By doing so, the study contributes with enhanced understanding of instructional quality in Swedish schools and a new angle from which to assess the need for future educational research and instructional development. The study draws on video recordings of 274 lessons gathered from 73 separate 7th grade classrooms of language arts and mathematics.

**Patterns of Instruction in Swedish Classrooms**

While teaching quality may be a culturally determined concept (Clarke et al., 2006), dimensions such as classroom organization, emotional support, instructional scaffolding, and cognitive, or intellectual, activation of students are frequently stressed in different educational traditions (Gitomer et al., 2014; Grossman et al., 2013). It has been reported that students in Nordic classrooms by comparison receive ample opportunities for dialogue and participation, although the interaction is often characterized as procedural rather than content-related or cognitively demanding (Klette et al., 2018). Teacher training and curricula in the Nordic countries is said to emphasize individualization and self-regulation among students and advocate less teacher transmission (Carlgren et al., 2006; Furuhagen et al., 2019). As recent research has pointed to “structured teaching” as an indicator of quality, it has been suggested that declining results for Sweden in international assessments, besides social and cultural factors, might be related to a less explicit, or “toned-down”, teacher-role in Swedish schools (Håkansson, 2015). Structured teaching is not a well-defined concept, but usually refers to an organized knowledge structure, scaffolding of students’ learning, well-organized time use, and engaged student participation in well-chosen tasks. To what extent such qualities characterize teaching in Swedish classrooms or not is less known, but some of the mentioned features are examined in the present study.

To define teaching quality necessarily needs to take subject-specific traditions into account. It has been suggested that Swedish language arts instruction is to a large extent skills-oriented (Bergman, 2007), but provides few opportunities for students to work with texts in order to build comprehension, and that explicit scaffolding practices of comprehension or production of texts are rare (Swedish Schools Inspectorate, 2012; Tengberg, 2019). Text discussions, when provided, allow students to express their thoughts and ideas about texts, but rarely challenge students’ understanding (Swedish Schools Inspectorate, 2012). In a similar vein, it has been observed that Swedish mathematics instruction has a strong focus on the development of procedural competence rather than problem-solving or reasoning competencies (Boesen et al., 2014; Swedish School Inspectorate, 2009). Teachers’ role in mathematics has been found to be passive “when it comes to specifying lesson goals, monitoring group work and orchestrating classroom discussions” (Van Steenbrugge, 2019).
& Ryve, 2018, p. 803), and the instruction judged not to be sufficiently varied to meet the different needs and capabilities of students (Swedish School Inspectorate, 2009). Plenary introduction followed by silent individual seatwork in textbooks appears to be the dominant instructional format in mathematics classrooms. Reichenberg (2018), who examined teaching practices in Swedish middle and lower secondary school, noted that more teacher-centered forms of instruction were found in social science, whereas group work was a less prevalent instructional format in all subjects.

In comparing the different modes of thought that underpin learning and instruction in mathematics and language arts, Kleve and Penne (2012) found, based on empirical data from classroom research, that whereas the aim of mathematics is to generalize and decontextualize knowledge to be able to explain mathematical phenomena, teaching and learning in language arts aims at contextualizing general knowledge to be able to interpret and make meaning of knowledge in specific contexts. Such differences of subject-specific ideals and epistemological traditions must be taken under considerations when analyzing quality of instruction in the two subjects.

Traditionally in Sweden, curricula, teacher education programs, and teacher guides contain few lesson plans or detailed instructions regarding lesson planning, thus placing a great deal of authority and professional independence on the teachers. Teacher beliefs about curricular content is therefore of vital importance. A comparative study of Finnish and Swedish teacher educators’ conceptualizations of effective teaching in mathematics showed that while Finnish teacher educators stressed the importance of clear presentations of mathematical content, homework, and specifying lesson goals, Swedish teacher educators emphasized to build on students’ ideas and connect content to everyday situations (Hemmi & Ryve, 2015). Such findings may be related to the results from classroom observations mentioned above, and to the suggestions made by Håkansson (2015) that a more structured leadership, perhaps a more traditional teacher-role, is rare in Swedish classrooms. However, the current knowledge of instructional quality in Swedish classrooms rely mainly on several small-scale studies that are promising and useful for designing hypotheses, but at the same time hard to relate to trends in large-scale assessments.

Observing Teaching Quality

In large-scale studies, teaching quality refers to features of instruction found important to students’ learning and is commonly understood as a multidimensional construct (Nilsen & Gustafsson, 2016). Features usually considered include the detail, clarity, and explicitness of instruction (Seidel & Shavelson, 2007), the classroom management required to make efficient use of lesson time and create an opportunity for learning (Praetorius et al., 2018), cognitive activation to encourage higher-level thinking among students (Baumert et al., 2010), and a supportive climate to motivate students, differentiate instruction, and provide functional guidance of students’ learning (Hattie, 2009; Praetorius et al., 2018). By turning such features of practice into observation variables, or rubrics, it is assumed that high-quality instruction may be identified by its relationship to measures of student achievement over time (Cohen & Goldhaber, 2016; Hanushek & Rivkin, 2010).

In the present study, we use the Protocol for Language Arts Teaching Observation (PLATO) as an observational lens to analyze classroom instruction. PLATO was developed by Grossman et al. (2013) to capture the quality of English language arts (ELA) instruction. According to Bell et al. (2019, p. 18), PLATO “privileges socioconstructivist approaches to learning but combines this with more cognitive approaches.” The composition of features captured in PLATO builds on research on practices proven critical to student learning in ELA, including strategy instruction (Palincsar & Brown, 1984), dialogic discourse (Nystrand et al., 1997), and cognitively activating tasks (Lipowsky et al., 2009). The instrument is organized into four broader domains: Instructional Scaffolding, Disciplinary Demand, Representations and Use of Content, and Classroom Environment, which are in turn divided into sub-elements, 12 elements in total (see Table 1). Recorded lessons are divided into 15-minute segments and scored on each element using a four-point scale. The theoretical assumption underlying its use in previous observation studies is that a high teacher average on
PLATO will predict a higher student achievement average. In the Measures of Effective Teaching (MET) study, PLATO demonstrated a stronger association with achievement in language arts than comparable observation instruments (Kane & Staiger, 2012) and was found to predict more ambitious student outcomes, such as critical thinking and reasoning better than low-inference comprehension measures (Grossman et al., 2014). In addition to the 12 elements presented in Table 1, studies using PLATO often include observations of overall instructional formats (whole class, group work, pair work, or individual seat work).

Although initially developed for ELA instruction, PLATO has also been used in mathematics (Cohen, 2018) and science (Kloser, 2014). Cohen (2018) examined teaching practices in American middle grade using PLATO and found significantly more explicit instruction in mathematics than in language arts. Preliminary results from an observation study of Norwegian mathematics and language arts in 8th grade (correspondent to Swedish 7th grade) found similar results in that high-quality strategy instruction and modeling were more common in mathematics than in language arts (Klette et al., 2017). The same study found that the highest PLATO scores were found in the domain Classroom Environment, indicating that disciplinary problem were highly infrequent. A study of Norwegian English L2 showed that while strategy use occurred frequently, explicit instruction of strategies was very scarce (Brevik, 2019).

**Aim and Research Questions**

In the present study, we use PLATO to analyze both language arts and mathematics instruction. The inclusion of the two subjects was not made primarily for comparative purpose, yet guided by previous research, we assumed that the different teaching traditions might cause variation in the results, and we therefore treat the two subjects separately. The study aims to provide empirically-based knowledge of prevalent patterns of instruction in Swedish lower secondary language arts and mathematics. With “patterns of instruction”, we refer both to instructional formats and research-based critical features of instructional quality. By doing so, the study wishes to contribute to enhanced understanding of instructional quality in Swedish schools and a new angle from which to assess the need for future educational research and instructional development. More specifically, this research addresses the following research questions:

1) To what extent are different instructional formats used by teachers in Swedish 7th grade language arts and mathematics?
2) To what extent are features of instructional quality present in Swedish 7th grade language arts and mathematics?

| Table 1. Domains and Elements of PLATO. |
|----------------------------------------|
| **Domain** | **Element** |
| Instructional Scaffolding | Modeling (MOD) |
| | Strategy Use and Instruction (SUI) |
| | Feedback (FDBK) |
| | Accommodations for Language Learning (ALL) |
| Disciplinary Demand | Intellectual Challenge (IC) |
| | Classroom Discourse (CD) |
| | Text-Based Instruction (TBI) |
| Representations and Use of Content | Representation of Content (ROC) |
| | Connections to Prior Knowledge (CPK) |
| | Purpose (PUR) |
| Classroom Environment | Behavior Management (BM) |
| | Time Management (TM) |

---

PLATO will predict a higher student achievement average. In the Measures of Effective Teaching (MET) study, PLATO demonstrated a stronger association with achievement in language arts than comparable observation instruments (Kane & Staiger, 2012) and was found to predict more ambitious student outcomes, such as critical thinking and reasoning better than low-inference comprehension measures (Grossman et al., 2014). In addition to the 12 elements presented in Table 1, studies using PLATO often include observations of overall instructional formats (whole class, group work, pair work, or individual seat work).

Although initially developed for ELA instruction, PLATO has also been used in mathematics (Cohen, 2018) and science (Kloser, 2014). Cohen (2018) examined teaching practices in American middle grade using PLATO and found significantly more explicit instruction in mathematics than in language arts. Preliminary results from an observation study of Norwegian mathematics and language arts in 8th grade (correspondent to Swedish 7th grade) found similar results in that high-quality strategy instruction and modeling were more common in mathematics than in language arts (Klette et al., 2017). The same study found that the highest PLATO scores were found in the domain Classroom Environment, indicating that disciplinary problem were highly infrequent. A study of Norwegian English L2 showed that while strategy use occurred frequently, explicit instruction of strategies was very scarce (Brevik, 2019).

**Aim and Research Questions**

In the present study, we use PLATO to analyze both language arts and mathematics instruction. The inclusion of the two subjects was not made primarily for comparative purpose, yet guided by previous research, we assumed that the different teaching traditions might cause variation in the results, and we therefore treat the two subjects separately. The study aims to provide empirically-based knowledge of prevalent patterns of instruction in Swedish lower secondary language arts and mathematics. With “patterns of instruction”, we refer both to instructional formats and research-based critical features of instructional quality. By doing so, the study wishes to contribute to enhanced understanding of instructional quality in Swedish schools and a new angle from which to assess the need for future educational research and instructional development. More specifically, this research addresses the following research questions:

1) To what extent are different instructional formats used by teachers in Swedish 7th grade language arts and mathematics?
2) To what extent are features of instructional quality present in Swedish 7th grade language arts and mathematics?
3) How do features of instructional quality differ between individual teachers in language arts and mathematics?

4) In what way are teacher characteristics (age, gender, teaching experience, and subject-specific qualifications) related to features of instructional quality?

**Methods**

**Study Design**

The study reproduced a research design employed by Klette et al. (2017) in Norway, in which lessons of Norwegian language arts and mathematics were video-recorded and analyzed using PLATO. Video-recordings were supplemented by photos from the classrooms, and statistics from coding of instructional formats and PLATO elements was used to answer the four research questions. To answer the fourth research question, data of teacher characteristics, such as age, gender, length of teaching experience, and extent of education, was also collected.

**Sampling Procedures and Participants**

The sample included 73 classrooms (38 in language arts and 35 in mathematics) distributed across 15 schools in nine municipalities around the country. The majority of the participating classes were observed in both language arts and mathematics. Schools were stratified to include a variation on the following variables: locality, immigrant proportion, students’ socioeconomic status, ownership form, grade averages and national test averages, and to match the national averages and distributions on these variables. However, schools were not randomized within strata.

As seen in Table 2, the sample is distributed over different types of communities and school forms. It is highly representative of the population in terms of academic achievement levels; however, the proportion of students with an immigrant background and the parent education level are

| School no | School locality | Immigr. proprot. | Higher ed. proprot. | Public/Charter | Merit Aver. | NT Aver. LA | NT Aver. MA |
|-----------|----------------|------------------|--------------------|---------------|-------------|-------------|-------------|
| 1         | Town           | 0.16             | 0.64               | Public        | 216.5       | 14.3        | 10.6        |
| 2         | Village        | 0.07             | 0.37               | Public        | 198.1       | 10.8        | 10.1        |
| 3         | Town           | 0.08             | 0.71               | Public        | 225.3       | 13.9        | 10.8        |
| 4         | Town           | 0.08             | 0.73               | Public        | 229.6       | 13.6        | 11.7        |
| 5         | Town           | 0.69             | 0.37               | Public        | 187.8       | 12.5        | 9.2         |
| 6         | Small town     | 0.14             | 0.46               | Public        | 205.9       | 12.5        | 10.2        |
| 7         | Village        | 0.29             | 0.31               | Public        | 217.2       | 13.1        | 12.4        |
| 8         | Town           | 0.14             | 0.67               | Public        | 242.1       | 15.3        | 12.4        |
| 9         | Town           | 0.2              | 0.53               | Public        | 201.5       | 13.4        | 10.3        |
| 10        | Town           | 0.19             | 0.54               | Public        | 201.0       | 12.4        | 9.0         |
| 11        | Small town     | 0.37             | 0.25               | Public        | 182.4       | 12.3        | 9.7         |
| 12        | City           | 0.1              | 0.87               | Charter       | 271.6       | 16.0        | 14.6        |
| 13        | Town           | 0.14             | 0.66               | Public        | 231.9       | 13.6        | —           |
| 14        | Town           | 0.23             | 0.51               | Public        | 211.4       | 14.0        | 11.4        |
| 15        | Town           | 0.29             | 0.46               | Public        | 203.0       | 12.7        | 11.0        |
| Sample average | 0.21 | 0.54 | 0.11 | 215.0 | 13.4 | 10.9 |
| National average | 0.25 | 0.59 | 0.20 | 216.9 | 13.6 | 11.2 |

*In accordance with definitions in the PISA School Questionnaire (Item SC001) (OECD, 2017). Proportion of students born abroad or both parents born abroad, data from the year of our data collection. Proportion of students with at least one parent with tertiary education in the year of data collection. Proportion of students with at least one parent with tertiary education in the year of data collection. Public or charter school. Average represents proportion of students in school year 7 in sample vs. in national population attending charter schools during the year of data collection. Based on student grades in all subjects in 9th grade. Average of the 5 years preceding the data collection. National test averages in 9th grade language arts and mathematics respectively. Scale from 0–20. Average of the 5 years preceding the data collection.
both somewhat low compared to the national averages. Moreover, the proportion of students in charter schools is low compared to the national average.

All schools were contacted through their principals, who forwarded our request to their teachers. All teachers and students were informed about the purpose of the research, the uses of the data, and their rights as participants in the research project. They were also given the opportunity to ask questions before the data collection began. Informed consent was signed by all participating teachers, students, and student guardians. Students who chose not to participate were seated in such a way that they could participate in lessons without being captured on video (see example of camera set-up in Figure 1, where the red seats were reserved for non-participating students). The fact that teachers self-selected to participate may result in bias, e.g., by including teachers who are generally more comfortable with their own teaching, and thus perhaps better than average. However, only two of the 17 schools initially approached declined participation, and on a majority of the 15 schools included, all teachers in the relevant age cohort agreed to participate, Thus the room for selection bias is judged to be limited in this is study.

The teachers finally included in the sample ranged in age, gender, and qualifications (see Table 3). They were generally well-experienced teachers (Mean = 16.2 years of practice, SD = 8.3). Further, the type and number of professional development courses attended ranged from none at all to several different courses. The sample thus provides a fair and diverse representation of language arts and mathematics teachers in Sweden. This does not necessarily mean that the teachers included in the sample represent the range of variation of teaching practices that may be found within the population of lower secondary language arts and mathematics teachers across the country. Results from the present study should therefore be treated as indicatory and not as grounds for generalization.

**Observations of Teaching**

Video observations were conducted using a set-up with two cameras (one in front to capture the students and one in the back to capture the teacher) and two microphones (one in the ceiling in

![Figure 1. Camera Set-up for Data Collection (picture: van Bommel).](image-url)
the middle of the room to capture student talk and one worn by the teacher to capture teacher talk), as shown in Figure 1. While teachers’ instruction is the focus of the study, students’ participation in terms of questions, comments, and other reactions are relevant to make valid inferences about teacher actions. The number and sample of lessons necessary to obtain a reliable representation of teachers’ teaching quality is a contested issue in recent research (Cohen & Goldhaber, 2016). Some studies indicate that four, even three, lessons may be sufficient to capture enough variation of individual teachers’ practice to enable reliable observation scores (Kane & Staiger, 2012). Kor (2011) suggests that to obtain a reliability of .80 on PLATO, at least five 15-minute segments of instruction should be observed. In the present study, each teacher was observed in three or four lessons, yielding an average of 12.7 segments per teacher. While the duration of lessons varied between schools, the minimum number of observations of a single teacher was seven 15-minute segments.

In order to provide a denser picture of the instruction captured on video, all observations were supplemented by photos of study material, white board annotations, classroom props, and student products. In line with previous research using PLATO (Grossman et al., 2013; Klette et al., 2017), these data were factored into the analysis of instructional quality.

Analytic Procedures

Recorded lessons (NLA = 141, NMA = 133) were divided into 15-minute segments (NLA = 458, NMA = 403) and scored using PLATO for instructional quality, and instructional format and content coverage. 15-minute segments are thus the case unit of statistical analysis in this study. All raters were trained and certified to use PLATO (a minimum requirement for a rater to become certified was to score at a reliability of 80%). To monitor reliability, regular meetings were held to score videos jointly and decide on critical issues and scoring rules. About 40% of all lesson segments were scored by two raters and disagreements settled through discussion.

Scoring by PLATO relies on qualitative criteria on each step of the four-point scale for each of the twelve elements. A general rule, however, is that low scores (1–2) are applied when the segment contains “almost no evidence” (1) or “limited evidence” (2) for the particular feature of interest. Conversely, high scores (3–4) are applied when the segment contains “evidence with some weaknesses” (3) or “consistent strong evidence” (4) of the feature of interest (Grossman, 2015). For example, the rubric description and scoring criteria for the element Strategy Use and Instruction is provided in Figure 2.

In addition to PLATO scoring, segments were coded for use of instructional format, i.e., whether whole-class, group work, pair work, or individual seatwork respectively were either present or not present, and which of them were the dominant format during each segment. The coding was based on the format directed by the teacher. Thus, if the teacher told students to work independently at their desks but some students co-operated in pairs, we counted this as individual seatwork. Conversely, if teachers instructed students to work in groups and many students made limited effort to cooperate with peers and mostly worked independently, we still counted this as group work, as it was the format provided by the teacher. As there was strong association between measures of presence and dominance of formats, only the statistics of dominant instructional formats are presented.
In the analysis, descriptive statistics on instructional format are presented both separately and related to statistics from PLATO scoring, while PLATO statistics is in turn presented both separately and as dependent variable of teacher variables such as age, gender, experience, and qualifications. It should be noted that while the sample is diverse, categorization on teacher variables generate small units of analysis and even though some differences are statistically significant, the results should be treated as indicators of which further corroboration is necessary.

Content Coverage

The observed instruction covered a range of content areas, although reading literature and writing texts of different kinds comprised a substantial part of the language arts lessons, while geometry and percentages & fractions were major areas in the observed mathematics instruction (see Table 4). In mathematics, the dominance of two areas probably resulted from the timing of observation, as many textbooks follow a similar planning throughout the year.

Results

The Use of Instructional Formats

As displayed in Figure 3, whole-class instruction and individual seatwork were the two dominant instructional formats in the observed classrooms. Whole-class instruction was the most frequently

Table 4. Content Coverage Observed in Language Arts and Mathematics.

| Area                                    | Language arts (total 458 segments) | Mathematics (total 403 segments) |
|-----------------------------------------|------------------------------------|----------------------------------|
| Present in segm. (N)                   |                                   | Present in segm. (N)             |
| Reading literature                      | 254                                | Algebra                          | 64 |
| Reading facts                          | 66                                 | Geometry                         | 134 |
| Literature discussions                 | 134                                | Arithmetic                       | 36  |
| Fact text discussions                  | 47                                 | Statistics                       | 10  |
| Text production                        | 231                                | Functions                        | 27  |
| Grammar/linguistics                    | 76                                 | Measures                         | 25  |
| Oral production                        | 7                                  | Percentages & fractions          | 148 |
| Listening                              | 61                                 | Problem solving                  | 64  |
| Culture/history of language or literature | 121                                | Other                            | 0   |

*Each segment could be coded for several content areas.*
(44.8%) used format in language arts, while individual seatwork covered a majority (51.6%) of lesson segments in mathematics. Pair work and group work made up a comparatively limited amount of the instruction in both subjects. Group work was more common in language arts than in mathematics, while pair work was somewhat more common in mathematics. A chi-square test of independence showed a significant association between subject and instructional format, $\chi^2(3, N = 860) = 45.45, p < 0.001$.

**Features of Instructional Quality as Defined by PLATO**

Table 5 presents mean PLATO scores for all separate elements and $p$-values for the difference between means in the two subjects, while Figure 4 displays distribution of scores across the four-point scale for the separate elements. According to the results, mathematics instruction appeared to be more explicit than language arts instruction and was more engaged with formal

**Table 5.** Means, Standard Deviations, and Ranges Across all PLATO Elements.

|              | MOD  | SUI  | FB   | ALL  | IC   | CD   | TBI  | ROC  | CPK  | PUR  | BM  | TM  | PLATO mean  |
|--------------|------|------|------|------|------|------|------|------|------|------|-----|-----|-------------|
| Language arts|      |      |      |      |      |      |      |      |      |      |     |     |             |
| Mean (N=458) | 1.52 | 1.62 | 2.03 | 1.66 | 2.27 | 1.66 | 2.27 | 2.40 | 1.57 | 2.13 | 3.69 | 3.47 | 2.18        |
| SD           | 0.75 | 0.83 | 0.67 | 0.73 | 0.71 | 0.73 | 1.02 | 0.68 | 0.75 | 0.38 | 0.66 | 0.78 | 0.32        |
| Indiv. mean high| 2.50 | 2.71 | 2.63 | 2.38 | 3.00 | 2.69 | 3.42 | 3.62 | 2.23 | 2.56 | 4.00 | 4.00 | 2.81        |
| Indiv. mean low| 1.00 | 1.00 | 1.50 | 1.00 | 1.43 | 1.00 | 1.00 | 1.90 | 1.08 | 1.92 | 1.70 | 1.54 | 1.75        |
| Mathematics  |      |      |      |      |      |      |      |      |      |      |     |     |             |
| Mean (N=403) | 2.17 | 2.54 | 2.21 | 2.03 | 2.50 | 1.74 | —    | 2.73 | 1.65 | 2.11 | 3.60 | 3.33 | 2.42        |
| SD           | 0.93 | 0.95 | 0.68 | 0.85 | 0.61 | 0.86 | —    | 0.83 | 0.76 | 0.38 | 0.70 | 0.86 | 0.43        |
| Indiv. mean high| 3.20 | 3.80 | 3.00 | 3.71 | 3.27 | 2.45 | —    | 3.70 | 2.38 | 2.50 | 4.00 | 4.00 | 2.84        |
| Indiv. mean low| 1.42 | 1.29 | 1.33 | 1.08 | 2.00 | 1.00 | —    | 2.00 | 1.14 | 1.57 | 2.27 | 2.64 | 1.95        |
| $p$-value of difference between means | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.14 | —    | 0.01 | 0.14 | 0.44 | 0.55 | 0.12 | 0.01        |

$^a$Individual mean high refers to the highest individual mean score by a single teacher in the sample. $^b$Individual mean low refers to the lowest individual mean score by a single teacher in the sample.
demonstration of how to solve specific tasks or how to understand subject-related problems. On average (and for a majority of the separate elements), teachers of mathematics scored significantly higher than teachers of language arts. Differences were particularly large for practices used to scaffold students’ learning, such as Modeling (MOD), Strategy Use and Instruction (SUI), Feedback (FB), and Accommodations for Language Learning (ALL), suggesting a pattern of distinction between teaching in the two subjects. Interestingly, in language arts, high scores (3–4) on SUI were more commonly found during whole-class instruction (58.0%) than during individual seatwork (15.9%), whereas in mathematics, high SUI scores were more common during seatwork (48.4%) than whole-class instruction (39.6%).

Average scores for Classroom Discourse (CD) and Connections to Prior Knowledge (CPK) were low compared to other variables in both subjects, suggesting that the opportunities for and quality of subject-related discussions are limited in both subjects, and that explicit connection between prior lessons and new content is rare. As shown in Figure 4, a high proportion of segments score 1–2 on CPK, suggesting that none or only vague connections are made. However, since explicit connections to prior lessons are likely to be concentrated to the beginning of a lesson rather than distributed evenly across it, we ran a separate analysis including only the first segment of each lesson. Results showed that 37.2% of all the first segments in language arts and 29.9% in mathematics scored at least a 3 on CPK. Thus, while explicit references to background knowledge or prior lessons that enable understanding of new content were rare in general across lessons, they occurred in about a third of all the lesson beginnings.

In contrast, scores on Representation and Use of Content (ROC) were comparatively high, indicating that accurate explanations and focus on conceptual understanding of subject content were regularly observed in both subjects. In both subjects, teachers also scored substantially higher on the Classroom Environment variables than on variables in the other three domains. A possible

Figure 4. Distribution of PLATO Scores across all Segments.

TBI is not included for comparability. Since TBI focuses on the use of text, it is clearly subject-specific for language arts and was commonly scored 1 in mathematics.
explanation for the difference is that the scoring of Behavior Management (BM) and Time Management (TM) takes a score of 4 as departure and uses evidence of various types of disruptions to adjust down the score, whereas scoring of the other ten elements takes 1 as a departure and uses observable qualities of teaching to adjust the score up. Nevertheless, high average scores on BM suggest that classrooms were mostly orderly and that instances of disruptions, distracting one or a few students from learning, were few. However, it does not categorically mean that students were always working concentratedly. The element does not capture, for instance, the rate at which students are unfocused or, in other ways, occupied by off-task behavior. Conversely, a score of 3 or lower on TM suggests that at least 2 of the 15 min in the segment are characterized by “downtime,” i.e., time spent on something other than subject-relevant work. This was evident in almost 40% of the observed segments in language arts and almost 50% of those in mathematics (see Figure 4).

**Differences of Instructional Quality between Teachers**

The distribution of PLATO scores between teachers in different classrooms shows that for the separate elements, as well as on average, teachers varied substantially in their teaching. Standard deviations (SD) for individual elements were generally large, and teachers’ individual mean scores ranged from 1.00 to over 2.50 for several elements (see Table 5). In language arts, the differences were notable, and teachers varied considerably in critical domains such as the amount of strategy instruction (SUI) provided, the level of intellectual challenge (IC), and the opportunities given to students for qualified use and production of texts (TBI). Differences between mathematics teachers were comparatively smaller but still substantial for several elements. In both subjects, the element Purpose (PUR) is an exception, for which SD was comparatively small, and a majority of teachers scored close to 2.00 on average.

Teachers thus varied substantially in terms of their employment of practices considered critical for students’ learning. Figures 5 and 6 show variation of average PLATO scores between observed teachers in both subjects. Although spread over more than ten recorded segments in three or four lessons and cut over twelve different instructional features, the scores vary considerably, suggesting that the quality of instruction that students received differed systematically between classrooms. From the point of view of PLATO, and in conjunction with the empirical studies validating its components by the power to predict achievement, teachers in different classrooms can be assumed to contribute differently to their students’ learning over time.

**Relationship between Teacher Characteristics and Instructional Quality**

Table 6 shows PLATO means for teachers with reference to gender, age, qualification in subject, and teaching experience. Female teachers scored significantly higher than male teachers in language.
arts, but not in mathematics. Age was a significant factor in both subjects, but with different results, when teachers were divided into three groups. In language arts, teachers above 50 years scored significantly higher than their younger colleagues, whereas in mathematics teachers aged 40–49 scored significantly higher than both younger and older teachers. Moreover, teachers with less than 90 ECTS in the subject scored significantly higher on PLATO in language arts, whereas no similar difference could be found in mathematics. Experience was not a significant factor in either of the subjects when comparing the most experienced half of the teachers to the less experienced half. It may be argued that experience might have shown an effect if instead we had compared the fairly unexperienced (e.g., less than five years of practice) to the more experienced (more than five years). However, this was not possible as very few teachers in the present dataset had less than even ten years of practice. Still, these findings may provide provisional clues to the systematic differences of teaching quality observed in different classrooms. However, although several of the differences presented in Table 6 are statistically significant, they remain small and the results should be treated with caution since the number of teachers counted in each cell are few.

Discussion

This study set out to investigate patterns of instruction in Swedish lower secondary language arts and mathematics. The findings of the study show that whole-class and individual seatwork activities were the dominant instructional formats. Mathematics instruction appeared to be more explicit than language arts instruction, marked, for instance, by a higher rate of scaffolding activities such as MOD, SUI, and FB, and by a higher PLATO mean score. In both subjects, however, the quality of instruction according to PLATO differed considerably between classrooms. Some of these differences were significantly related to teachers’ gender, age, and length of education, but the relationships were inconsistent between subjects, and the study does not provide a clear explanation to these findings.

The findings concerning instructional format are in line with those of previous studies (Reichenberg, 2018; Swedish Schools Inspectorate, 2012; Van Steenbrugge & Ryve, 2018), though somewhat divergent from commonly held assumptions about a student-centered and participatory tradition in Swedish compulsory school. While students, particularly in mathematics, spent a large amount of time on individual task completion, pair and group work occupied minor shares of the instructional time in both subjects, and teacher-led whole-class instruction was pervasive, particularly in language arts. However, to determine that instructional activities are teacher-led reveals in itself little about their potential instructional qualities. The analysis does not expose the extent to which whole-class instruction was devoted to teacher lecturing, watching videos, discussing, or solving problems on the whiteboard together. In fact, the instruction we observed included all of those activities, but to determine whether teacher-led whole-class instruction
Table 6. PLATO Means Related to Teacher Characteristics.

|                | Gender | Age     | Qualification in subject (ECTS)<sup>a</sup> | Experience of teaching |
|----------------|--------|---------|---------------------------------------------|------------------------|
|                | Female | 20–39   | 40–49 | 50+ | 0–60/90 | 61/91+ | 0–14 years | 15 years or more |
| Language arts  | mean   | 2.13**  | 2.06  | 2.09 | 2.19** | 2.13** | 2.05       | 2.09  | 2.11       |
|                | SD     | 0.30    | 0.32  | 0.28 | 0.33   | 0.29   | 0.32       | 0.30  | 0.32       |
| Mathematics    | mean   | 2.31    | 2.29  | 2.28 | 2.38*  | 2.24   | 2.31       | 2.30  | 2.31       |
|                | SD     | 0.37    | 0.41  | 0.40 | 0.38   | 0.41   | 0.36       | 0.39  | 0.39       |

* mean significantly higher at p < .05, ** mean significantly higher at p < .01.<sup>a</sup>Teachers of language arts were divided at 0–90 or 90+ ECTS, while in math they were divided at 0–60 or 61+, since only one out of 35 math teachers had more than 90 ECTS compared to 14 out of 38 language arts teachers (see Table 3 above).
contained more explicit scaffolding and teacher direction, and that it is therefore less student-centered or less focused on cooperation, requires further analysis. It is worth noting, however, that high-quality strategy instruction SUI scores 3–4) occurred more frequently during individual work in mathematics than during whole-class instruction. Thus, there is no clear evidence in the study for incongruity between structured teacher guidance and individual work.

Several findings are consistent with those of similar studies using PLATO conducted in American and Norwegian schools in similar grade levels. The MET Study also reported low levels of scaffolding activities, such as MOD and SUI, and few behavioral problems in the classrooms (Grossman et al., 2013; Kane & Staiger, 2012). Similarly, Norwegian language arts and mathematics teachers both scored low on scaffolding activities such as CPK, FB, and SUI (Klette et al., 2017), though explicit instructional elements such as SUI and MOD where more frequent during mathematics lessons than language arts lessons. In contrast, presenting learning goals (PUR) and connecting new material to prior academic knowledge (CPK) were features more frequently found in American classrooms than in the Swedish classrooms in this study.5

Thereby, the results of the study provide a useful and up-to-date overview of the nature and quality of teaching in Swedish lower secondary language arts and mathematics, an overview that may inform educational debate and serve as a point of reference for future research for comparative purposes and for mapping trends of instructional change. In addition to the light cast on teachers’ scaffolding activities, the study directs attention to the low level of high-quality and content-related classroom dialogue (low scores on CD), often referred to in previous research as an important feature of high-quality instruction (Murphy et al., 2009; Webb et al., 2019). Contrary to what is often assumed about Swedish classrooms, extended classroom dialogue appeared to be scarce according to the findings of this study. Similarly, the findings highlight the need for discussion about the cognitive activation and intellectual challenge offered by daily classroom activities. It seems clear that future efforts at instructional improvement should carefully consider how instruction might provide elaborated and extended opportunities for engagement with content beyond routinized assignments. Engaging students in higher-order thinking by asking them to analyze composite problems, combine perspectives, and provide justification for their answers is necessary for elaborated knowledge construction (Praetorius et al., 2018). How such activities may be thoughtfully planned for and orchestrated in the classroom is a key area to which future educational debate and improvement enterprises should commit.

**Strategy Instruction and Content Specificity**

By including observation of instruction in two subjects, the study design allows for an investigation of the content specificity of what may otherwise (somewhat superficially) be labeled as generic features of instruction. Against the wealth of prior research that has suggested the importance of strategy instruction, particularly in language arts (Mokhtari, 2017; Palincsar & Brown, 1984), it is interesting to note that strategy instruction was observed much more often in mathematics. PLATO defines strategy as “a flexible method (…) that can be applied [by the student] in multiple contexts” (Grossman, 2015). It is reasonable to ask, given the observed difference, whether the similar instances of teacher–student–content interaction were treated as strategy instruction in the two subjects. While this question may deserve subsequent analysis, it is fair to hypothesize that although teachers in both subjects probably strived to provide students with flexible methods that could be applied in multiple contexts, generic and formulaic step-by-step procedures that are quite common in mathematics were less naturally discerned in typical language arts content. This also aligns with the orientation towards generalization and decontextualization of rules in mathematics vs. the focus on interpretation *within* context in language arts, suggested by Kleve and Penne (2012). Thus,

5Note that data collection for the MET study was conducted in 2008–2010, while data for the present study were collected in 2017–2019.
explicit instruction of flexible methods is perhaps more naturally attached to teacher guidance in mathematics. This was also found by Cohen (2018) in a recent study of American middle grade instruction in which she distinguished Procedural Strategy Instruction (i.e., explicit teacher-led instruction of rules or principles without discussion of alternative approaches) from Conceptual Strategy Instruction (i.e., exploring through discussion and reasoning why and how strategies work). She found that while mathematics contained significantly more of the former type than language arts, the latter type was equally present in both subjects. Whether this distinction offers an explanation for the observed difference in the present study will require further analysis. Nevertheless, it highlights the need for careful consideration of the different forms of strategy instruction in different content areas and the limit to which concepts such as strategy instruction may be analytically useful without closer attention to the specific content being taught.

**Limitations and Future Research**

There are some limitations to this study that should be considered when interpreting the findings. First, by presenting distributions of PLATO means across classrooms, the study implies a theoretical assumption of a linear relationship between quality indicators and student learning, i.e., that higher PLATO scores are proportionally related to increased learning. While previous studies in the US have indeed indicated a relationship between PLATO mean scores and higher learning outcomes (Kane & Staiger, 2012), this remains to be validated in a Nordic educational context. Additionally, the exact nature of this relationship remains unclear. For instance, we do not know whether teacher impact on student achievement is primarily related to high PLATO averages in general or to discrete occurrences of high-end (level 4) performances on certain elements. While the predictive value of PLATO in the present data set will be investigated in a subsequent study, the statistical averages presented in this study do provide important empirical evidence of prevalent instructional patterns in Swedish classrooms. Thanks to the study, we now know significantly more than before about the actual teaching practices of lower secondary language arts and mathematics.

Second, it should be underscored that although the sample used for the present study is larger than that of most previous classroom observation studies conducted in Sweden, and was stratified to include a representative variation of relevant school variables, it is still too small, and not randomized within strata, to be representative of the national population of 7th grade classrooms. Thus, while the study provides an important contribution and a benchmark for future large-scale classroom research, the frequencies and distributions across variables cannot be generalized and should rather be treated as indications requiring further corroboration.

Finally, it should be emphasized that while we speak generally in the article of the object studied as lower secondary instruction rather than 7th grade instruction, or mid-year 7th grade instruction, some patterns of classroom practice may be more grade level-specific than others. We have no evidence to infer that the observed patterns extend beyond 7th grade, even in the schools that were studied, or that the observed teachers are themselves stable across the different classrooms in which they teach. Student maturity, motivation, and content complexity may affect instructional characteristics and constrain the possibility of generalizing observed patterns across grade levels. However, it is neither justified by evidence to assume that such extensions are impossible. It is highly likely, for instance, that the challenge that some teachers experience in providing adequate room for quality dialogue in the classroom, or the tendency of other teachers to deliver constructive and substantive, rather than vague and procedural, feedback on students’ work, are characteristics that cut across grade levels. Hence, even though statistical data cannot be generalized across grades, it reasonable to infer that the present study provides a relevant and interesting pointer to what may also occur in adjacent school years.
**Implications for Practice**

The most important finding of the study is perhaps that teachers in different classrooms vary substantially and systematically on high-inference variables known from previous research to be vital for supporting student achievement gains over time. In other words, the instruction in different classrooms, sometimes neighboring classrooms in the same school, is of systematically different quality. This offers a key implication for future planning of instructional improvement and professional development initiatives. It suggests that a one-size-fits-all model for professional development—for instance, that all teachers ought to increase the amount of explicit strategy instruction, develop their capacity for intellectually challenging dialogue, or more carefully mind the time-on-task ratio of their lessons—may be a less fruitful design because it does not respond to the individual development needs of teachers. A more efficient approach would be to tailor improvement efforts closely to individual teachers and their shifting ambitions and aptitudes. Such attempts need not contain personal programs for every single teacher; rather, they should be designed to respond to teacher performances and point to reasonable development trajectories. An ongoing example is found in collaborative practice-based research projects, where observation and feedback on relevant features of instruction are used to improve teaching practices (Allen et al., 2011; Gore et al., 2017; Tengberg & Wejrum, 2021). However, it should be kept in mind that identification of relevant variables for focused and systematized observation and feedback requires large-scale observation studies in authentic classrooms, both for mapping the terrain of current practice and for determining the potential impact of different features of instruction. For this reason, the results of the present study may provide an important contribution to both the practice of and the research on educational improvement.

**Disclosure Statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by The Swedish Research Council: [Grant Number 2017-03544].

**ORCID**

Michael Tengberg [http://orcid.org/0000-0002-9375-9512](http://orcid.org/0000-0002-9375-9512)

Jorryt van Bommel [http://orcid.org/0000-0001-6525-9871](http://orcid.org/0000-0001-6525-9871)

Marie Nilsberth [http://orcid.org/0000-0002-7691-2490](http://orcid.org/0000-0002-7691-2490)

Michael Walkert [http://orcid.org/0000-0001-7653-9289](http://orcid.org/0000-0001-7653-9289)

Anna Nissen [http://orcid.org/0000-0001-7653-9289](http://orcid.org/0000-0001-7653-9289)

**References**

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034–1037. [https://doi.org/10.1126/science.1207998](https://doi.org/10.1126/science.1207998)

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers’ mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. [https://doi.org/10.3102/0028312009345157](https://doi.org/10.3102/0028312009345157)

Bell, C. A., Dobbelæer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 30(1), 3–29. [https://doi.org/10.1080/09243453.2018.1539014](https://doi.org/10.1080/09243453.2018.1539014)

Bergman, L. (2007). *Gymnasieskolans svenskämnen: En studie av svenskundervisningen i fyra gymnasieklasser*. Diss. Lund: Lunds universitet.

Blömeke, S., & Olsen, R. V. (2019). Consistency of results regarding teacher effects across subjects, school levels, outcomes and countries. *Teaching and Teacher Education*, 77, 170–182. [https://doi.org/10.1016/j.tate.2018.09.018](https://doi.org/10.1016/j.tate.2018.09.018)
Lipowsky, F., Rakocy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students’ understanding of the Pythagorean theorem. *Learning and Instruction, 19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Mokhtari, K. (2017). *Improving reading comprehension through metacognitive reading strategies instruction*. Rowman & Littlefield.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students’ comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101*(3), 740–764. https://doi.org/10.1037/a0015576

Nilsen, T. E., & Gustafsson, J. E. E. (2016). *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2). Springer International Publishing.

Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue*. Teachers College, Columbia University.

OECD. (2017). *School questionnaire for PISA 2018. Main survey version*. https://www.oecd.org/pisa/data/2018database/CY7_201710_QST_MS_SCQ_NoNotes_final.pdf

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM, 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Reichenberg, O. (2018). Identifying mechanisms of teaching practices: A study in Swedish comprehensive schooling. *Scandinavia Journal of Educational Research, 62*(2), 213–228. https://doi.org/10.1080/00313831.2016.1212262

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0031383107310317

Slater, H., Davies, N. M., & Burgess, S. (2012). Do teachers matter? Measuring the variation in teacher effectiveness in England. *Oxford Bulletin of Economics and Statistics, 74*(5), 629–645. https://doi.org/10.1111/j.1468-0084.2011.00666.x

Swedish Schools Inspectorate. (2009). Undervisningen i matematik – utbildningens innehåll och ändamålsenlighet. Kvalitetsgranskning. Rapport 2009:5. Swedish Schools Inspectorate.

Swedish Schools Inspectorate. (2012). Läsundervisning inom ämnet svenska för årskurs 7–9. Kvalitetsgranskning. Rapport 2012:10. Swedish Schools Inspectorate.

Tengberg, M. (2019). Textanvändning och texttolkning i svenskundervisningen på högstadiet. *Nordic Journal of Literacy Research, 5*(1), 18–37. https://doi.org/10.23865/njlr.v5.1488

Tengberg, M., & Wejrum, M. (2021). Observation och återkoppling med fokus på utvecklad undervisning: Professionsutveckling med hjälp av PLATO. *Acta Didactica Norden, 15*(1), 1–22. https://doi.org/10.5617/adno.7985

Van Steenbrugge, H., & Ryve, A. (2018). Developing a reform mathematics curriculum program in Sweden: Relating international research and the local context. *ZDM Mathematics Education, 50*(5), 801–812. https://doi.org/10.1007/s11888-018-0972-γ

Webb, N. M., Franke, M. L., Ing, M., Turrou, A. C., Johnson, N. C., & Zimmerman, J. (2019). Teacher practices that promote productive dialogue and learning in mathematics classrooms. *International Journal of Educational Research, 97*, 176–186. https://doi.org/10.1016/j.ijer.2017.07.009