

# Subject specificity matters? Investigating teaching quality across subjects and countries

Chair                    Kirsti Klette, University of Oslo  
Discussants:        Courtney Bell, Madison University

## Symposium Abstract

Significant advancements have been made in conceptualizing, operationalizing, and measuring teaching quality over the last two decades. The development of observation systems has paved the way for a more targeted and systematic measurement of features of teaching quality across contexts, recognized at the level of different observation manuals be it subject-specific manuals or generic manuals. In addition, and often as a supplement, student perception surveys have been used to provide additional information about the quality of teaching and instruction (van der Scheer, et al., 2019). The goal of this symposium is to present the results of an effort to study instruction across the Nordic countries, drawing on observation- and student survey measures from lower secondary classrooms in the three subjects: mathematics, language arts, and social science education.

A key aspect to be discussed is the role subject specific versus generic aspects of teaching quality. In the current study, we used the Protocol for Language Art Teaching Observation (PLATO) manual (Grossman et al., (2013) and the Tripod student perception survey (Ferguson, 2015) as generic measures of teaching quality – applying both instruments to all three subjects and lessons. Our analyses indicate. subject specific differences across subjects but also generic features of teaching quality that cut across all countries.

The papers within this symposium show how systematic coding, and student survey scores, can serve to generate broad pictures of instructional practices across subjects and classrooms, thus serve as a starting point for quantitative and qualitative explorations of specific subjects and lessons. The symposium demonstrates how: (i) a common framework and conceptual language provided by the observation manual can serve as the starting point for supporting collaborations in the study of teaching quality, (ii) how we used the observation scores (i.e. high scores) as sampling criteria to dig more closely into specific instructional practices such as features of Classroom Discourse and Intellectual Challenge within and across subjects countries, and (iii) how student perceptions (as measure by the Tripod Survey) might feed into and expand our understanding of teaching quality. The four papers provide a new step for our comparative classroom ambition in the Nordic Center of Excellence “Quality in Nordic Teaching” (QUINT). The first paper provides an overview of key findings based on using a common observational measure (the PLATO manual) across Nordic lower-secondary classrooms (n= 144) in the three subject areas. The second paper report from student survey perceptions (n= 2501) on teaching quality across subjects and countries. The third and fourth paper dig into, respectively, features of high quality classroom discourses and characteristics of cognitively activating classrooms in all three subjects and countries. Together these four papers summarize features of quality classroom teaching and learning and discuss how this might vary between subjects and countries.

## References:

- Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value added scores. *American Journal of Education*, 119(3), 445–470.
- Ferguson, R. F. (2015). *The influence of teaching beyond standardized test scores: Engagement, mindsets, and agency. A study of 16,000 sixth through ninth grade classrooms*. Report from The Achievement Gap Initiative. Harvard University.
- van der Scheer, E. A., Bijlsma, H. J. E. & Glas, C. A. W. (2019). Validity and reliability of student perceptions of teaching quality in primary education. *School Effectiveness and School Improvement*, 30(1), 30–50. <https://doi.org/10.1080/09243453.2018.1539015>

## **Paper 1: Observation manuals as lenses into classroom teaching – towards a common language of instruction?**

*Presenting Author: Kirsti Klette,*

*Co-Authors: Astrid, Roe, Marte Blikstad-Balas, Mark White, Michael Tengberg*

### **Abstract**

Classroom-based observational research with observation systems holds the potential to drive systematic and cumulative research that supports the comparison of instructional practices across studies and contexts (Bell et al., 2019). This paper presents the results of applying the Protocol for Language Arts Teaching Observation (PLATO) across Nordic lower secondary classrooms. Although initially developed for the purpose of measuring aspects of language arts classrooms, we also used PLATO to measure social science and mathematics instruction.

*Theoretical framework:* PLATO (Grossman et al., 2013) is tailored to assess instruction; encompassing 12 elements of instruction highlighted in existing literature on adolescent literacy as well as effective instruction in secondary language arts education. It is a systematic and validated protocol (Cohen & Grossman, 2016) that resonates well with key aspects of instructional quality (Kane & Staiger, 2012, Klieme et al., 2009; Nilsen & Gustafsson, 2016) as summarized in the research literature (e.g. instructional clarity, cognitive demand, discourse features and supportive climate).

*Methods and Data sources:* We report on video observations from respectively 52 language arts, 54 mathematics 8<sup>th</sup> grade classrooms and 38 Social Science classrooms, from which 470 lessons were observed. The classrooms were sampled to maximize the diversity across key criteria thought to impact school quality in the Nordic lower secondary context. The video recordings of lessons were coded using the PLATO observation tool. All raters underwent standard PLATO training and certification. Recorded lessons were divided into equal-interval 15-minute segments for coding (n=1380 segments). All segments were scored for all 12 elements, using a four-point scale where scores 1-2 are towards the low end and 3-4 are at the high end, and with sufficient interrater reliability agreement consistent with previous studies.

*Results:* Three main findings were revealed: First, PLATO seem to be able to capture some key patterns of instruction across the Nordic classrooms, however with a certain tendency to privilege explicit instruction. Second, PLATO scores across the sampled classrooms suggest similar broad patterns across countries in the three subjects. The observed teachers score consistently high on the elements Behavioral Management and Time Management, mediocre (towards low) on the elements Intellectual Challenge, Classroom Discourse, and Representation of Content. They score systematically low on Strategy Instruction, Modelling, and Feedback. There are however interesting variations both within and across classrooms and countries. Third, although designed for language arts instruction, the PLATO scores in mathematics are systematically higher in all five countries, especially for elements related to explicit instruction such as Modelling and Strategy Instruction. This shed light on issues of generic versus subject specific instruments when measuring teaching quality.

*Significance:* These findings nurture multiple conclusions and interpretations: PLATO provides reliable and qualified information about instructional practices in Nordic classrooms; gives an overview of specific patterns and strengths; works as a ‘diagnostic tool’ for more systematic work on targeted instructional elements (e.g. Feedback, Strategy Instruction, Classroom Discourse). As such, coding manuals serve multiple functions – for empirical validation of conceptions of teaching, as a diagnostic tool, and as a language for analyzing and describing features of teaching quality.

*References:*

- Bell, C. [Dobbelaer](#), M.J., Klette, K., Visscher, A (2019).Qualities of classroom observation systems *School Effectiveness and School Improvement*, 30 (7) 1–27.  
DOI:10.1080/09243453.2018.
- Cohen, J., & Grossman, P. (2016). Respecting complexity in measures of teaching: Keeping students and schools in focus. *Teaching and Teacher Education*, 55, 308-317.
- Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education*, 119(3), 445–470.
- Kane, T. J., & Staiger, D. O. (2012). Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains. Research Paper. MET Project. Bill & Melinda Gates Foundation.
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. I T. Janik & T. Seidel (eds.), *The power of video studies in investigating teaching and learning in the classroom* (137–160). New York: Waxmann Publishing Co.
- Nilsen, T., & Gustafsson, J.-E. (red.). (2016). *Teacher quality, instructional quality and student outcome. Relationships across countries, cohorts and time.* (Vol. 2). Springer.

## **Paper 2: Teaching quality in Nordic classrooms from the perception of students**

*Michael Tengberg, Berglind Gísladóttir, Astrid Roe & Anders Stig Christensen*

### **Abstract**

Knowing what works is critical to improving instruction. To the individual teacher, however, or to a single school, a community, or even a national education system, knowing one's particular strengths and weaknesses are equally important. Therefore, beyond identifying features of teaching that are effective for student learning, appropriate improvement of teaching, in a given system, also requires site-specific knowledge of prevalent classroom practices. For this purpose, comparative examination across educational contexts is essential in order to generate in-depth understanding of both strengths and development areas (Suter, 2019; Xu & Clarke, 2019). As supplement to classroom observations, evidence of instructional patterns and qualities is often gathered through student surveys. Recent development and validation of survey instruments suggests that student perceptions can provide reliable estimates of teacher performance (van der Scheer et al., 2019), and predict both achievement and affective outcomes (Fauth et al., 2019; Wallace et al., 2016).

In order to contribute with extended knowledge about prevalent characteristics of teaching in Nordic lower secondary classrooms, this presentation will report findings from a survey of student perceptions of teaching practices in Denmark, Finland, Iceland, Norway, and Sweden. Data was gathered from lower secondary language arts, mathematics, and social science classrooms in order to compare aspects of perceived teaching quality across subjects and countries. The following research questions guided the study:

- 1) How do lower secondary students in the Nordic countries perceive the teaching they receive on aspects of teaching quality?
- 2) To what extent do student perceptions of teaching quality vary between different subjects (language arts, mathematics, and social science)?
- 3) To what extent do student perceptions of teaching quality vary between classrooms?

The study drew on 2.501 responses to the Tripod survey (Ferguson, 2015) by Nordic students in lower secondary school (13–15 yrs). After confirmation of measurement reliability, including model fit of dimensional structure, descriptive statistical analyses on subscale level, including t-testing for statistical significance, was used to answer the three research questions.

The study identifies a range of interesting patterns of perceived teaching quality across subjects and countries. Among these patterns we find for example that lower secondary students in the Nordic countries report generally high regard for their teachers' capacity of clarifying and explaining content, and comparatively lower appreciation for their teachers' ability to make learning interesting and enjoyable, and for inviting students to share their ideas and speak their minds of the work done in class. The latter is especially pronounced in language arts and mathematics. Comparison between subjects also indicates that Nordic students generally perceive more teacher care and less academic challenge in language arts than in both mathematics and social science. In addition to general patterns, however, the study shows large amount of variation between classrooms within both countries and subjects. These results will be discussed with reference to prior research, and in relation to possible implications for practice.

*References:*

- Fauth, B., Decristan, J., Decker, A.-T., Büttner, G. Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality *Teaching and Teacher Education*, 86, 102882. <https://doi.org/10.1016/j.tate.2019.102882>
- Ferguson, R. F. (2015). *The influence of teaching beyond standardized test scores: Engagement, mindsets, and agency. A study of 16,000 sixth through ninth grade classrooms*. Report from The Achievement Gap Initiative. Harvard University.
- Suter, L. E. (2019). The status of comparative education research in the 21st century: An empiricist's views." In L. E. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 3–24). SAGE.
- van der Scheer, E. A., Bijlsma, H. J. E. & Glas, C. A. W. (2019). Validity and reliability of student perceptions of teaching quality in primary education. *School Effectiveness and School Improvement*, 30(1), 30–50. <https://doi.org/10.1080/09243453.2018.1539015>
- Wallace, T. L., Kelcey, B., & Ruzek, E. (2016). What can student perception surveys tell us about teaching? Empirically testing the underlying structure of the tripod student perception survey. *American Educational Research Journal*, 53(6), 1834–1868. <https://doi.org/10.3102/0002831216671864>
- Xu, L., & Clarke, D. (2019). Validity and comparability in cross-cultural video studies of classrooms In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education: Cross-disciplinary perspectives* (pp. 19–33). Routledge.

### **Paper 3: Features of Discourse in Nordic Classrooms**

*Presenting Author: Berglind Gísladóttir (with: Kirsti Klette, Camilla Gudmundsdóttir Magnusson, Alexander Selling, Anders Stig Christensen, Peter Aashamar, and Jennifer Luoto)*

#### **Abstract**

Classroom discourse plays an important role in student learning as it promotes critical thinking and deeper understanding by encouraging students to articulate their thoughts and engage with diverse perspectives. This interactive process not only enhances communication skills but also promotes active learning, as students are more deeply involved in the educational process through discussion and debate (Mercer et al., 2020, Resnick et al., 2015). Classroom discourse can transform passive learning into active exchange, allowing students to construct and refine their understanding through verbal expression of complex concepts. These interactions between students and teacher and among students promotes communication skills, as students are more actively involved in the educational process through discussion and debate (Howe & Abedin, 2013)

With the use of video data, the current study aims to provide some insight into the quality of classroom discourse in Nordic lower secondary classrooms. Video data was collected in Denmark, Finland, Iceland, Norway and Sweden in mathematics, language arts and social science. The analytical framework PLATO (Protocol for Language Arts Teaching Observation) was used to assess the quality of classroom discourse in the video data (Grossman, 2015). Recorded lessons were coded in 15-minute segments where both the opportunities given for student talk and the teacher's uptake of student responses were evaluated on a four-point scale. In our analysis we focused on lessons with high levels of teacher uptake and student participation, focusing on (a) the types of questions posed by the teacher (b) the nature of teacher uptake and (c) what characterized students' utterances.

Findings reveal a universal pattern across countries where discourse was typically driven by teacher questions followed by student answers, reflecting the traditional Initiation-Response-Feedback (IRF) pattern. We also found that student utterances were mostly brief responses to teacher questions even though the brevity was somewhat dependent on the type of question posed. Furthermore, a notable gap in our study was how infrequently students built upon each other's ideas.

Distinct variations were observed in question types across subjects. Teachers of Social Science used more open questions, encouraging more diverse viewpoints, while teachers in Mathematics predominantly used closed questions for factual recall. In Language Arts teachers showed a more balanced approach using a mix of open and close question. This suggests that the subject matter influences the type of discourse, with open questions in Social Science and Language Arts encouraging broader discussions, unlike the focus on procedural knowledge in Mathematics.

#### *References:*

- Grossman, P. (2015). Protocol for Language Arts Teaching Observations (PLATO 5.0). Stanford University.
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge journal of education*, 43(3), 325-356.
- Mercer, N., Wegerif, R. og Major, L. (ritstjórar). (2020). *The Routledge international handbook of research on dialogic education*. Routledge.
- Resnick, L. B., Asterhan, C. S. C. and Clarke, S. N. (Ed). (2015). *Socializing intelligence through academic talk and dialogue*. American Educational Research Association (AERA).

## **Paper 4: Intellectual challenge in Nordic classrooms**

*Jóhann Örn Sigurjónsson*

### **Abstract**

One central question for teaching observations is what the teacher expects students to do. Intellectual challenge concerns the selection and use of tasks and the richness and rigour of facilitated cognitive activity. From the perspective of the Three Basic Dimensions, it incorporates some key components of cognitive activation, which has been linked to both increased student achievement and enjoyment (Praetorius et al., 2018). The way teachers construct their lessons in this regard – i.e., to what extent assigned tasks require analysis and inferential thinking, and whether they are implemented to require students to justify and reason – impacts what opportunities are created for students to develop skills that are widely deemed critical for their education.

The current study aims to deliver insights into levels of intellectual challenge in Nordic lower secondary classrooms based on classroom video data from Denmark, Finland, Iceland, Norway and Sweden in three subjects: mathematics, language arts and social science. The Protocol for Language Arts Teaching Observations (PLATO) framework was used to score 425 lessons comprising the LISA Nordic video database (Grossman, 2015). In PLATO, each 15-minute segment is scored on a scale from 1 (low, weak evidence) to 4 (high, strong evidence). At the low-end of the intellectual challenge scale, students mostly engage in activities that require recall or rote thinking or are passive participants. At the high-end, students are mostly engaged in activities that require analytic or inferential thinking. An important component of intellectual challenge is that the score can be advanced or degraded by one point based on the nature of teacher questions and comments in relation to the challenge of the task as initially presented. The analysis considers maximum segment scores at the lesson level to privilege the presence of high-level intellectual challenge, being mindful that different parts of lessons may target different dimension of teaching quality. Selected lessons with 4-level segment scores were coded minute-by-minute for instructional formats.

The findings show that across subjects, a similar proportion of lessons scored consistently at the low-end, between 40-45%. Conversely, this means that in each subject, between 55-60% of lessons had at least one segment scored at the high-end. The proportion of lessons containing a 4-level segment was quite low: 9% in mathematics, 13% in language arts and 13% in social science. Only two countries had lessons at the 4-level in social science, but overall, the distribution of maximum segment scores within lessons was relatively similar between countries.

These findings suggest that intellectual challenge is a dimension of teaching that may need increased attention in a Nordic context, with possible implications for teacher education and professional development. Example vignettes from high-level lessons and instructional format timelines will be presented to spur discussions on similarities and differences to common teaching repertoires in a Nordic context and a dialogue for conceivable pathways forward in theory and practice.

### *References:*

- Grossman, P. (2015). Protocol for Language Arts Teaching Observations (PLATO 5.0). StanfordUniversity.
- 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>