The Emperor’s new clothes: PISA, TIMSS and Finnish mathematics

Paul Andrews
Stockholm University

Abstract
For nearly fifteen years, due to repeated successes on the Programme of International Student Assessment (PISA), Finnish education in general and mathematics education in particular have been construed internationally as benchmarks. In what is essentially a review paper I consider how the Finns explain their students’ repeated PISA successes before contrasting these explanations with observational evidence indicating that typical classroom practice is unlikely to account for such successes. In addition, I examine the relative failure of Finnish students on the Trends in International Mathematics and Science Study (TIMSS), particularly with respect to algebra and geometry, and highlight the extent to which Finnish students may be inadequately prepared for higher study of mathematics. I close by indicating that continued interest in Finland as a source of excellence in mathematics teaching may be misguided and that other European systems, like Flanders, may provide better warranted research locations for those interested in transferable insights.

Introduction
Finnish performance on the mathematics component of the Organisation for Economic Cooperation and Development’s (OECD) Programme of International Student Assessment (PISA) has created much international interest. Finnish results have been viewed both internally and externally, as exceptional, although the recent publication of the PISA 2012 results showed scores significantly lower than previously (OECD, 2013). However, prior to this, such has been the interest generated by Finnish successes that envoys from all around the world have visited Helsinki to uncover the story behind its success (Laukkanen, 2008). However, by drawing on available literature, I try to show not only how Finnish PISA performance may have been due to factors other than the quality of mathematics instruction but also why it is naïve to assume that success on any form of international test is a guarantee of transferable pedagogical quality.

So, by way of a starting point, let us examine the nature of PISA and its three year cycle since the first iteration in 2000. Its objective has been to evaluate the extent to which 15 year-old students “are prepared to meet the challenges of today’s societies” (OECD, 2003, p. 9). This it addresses by means of assessments of students’ literacy, mathematical literacy and scientific literacy at age 15, with each being the primary focus every third cycle beginning with literacy in 2000. With respect to mathematics, especially in 2003 and 2012, explicit attention has been paid to problems that move “beyond the kinds of situations and problems typically encountered in school classrooms” towards people’s daily lives and the sorts of problems that expect the application of mathematical skills in unfamiliar contexts, and which require “decisions about what knowledge may be relevant, and how it might usefully be applied” (OECD, 2003, p.24). In other words, PISA focuses “on the capacity of students to put mathematical knowledge into functional use in a multitude of different situations in varied, reflective and insight based ways” (Schleicher, 2007, p. 351). Such problems exemplify the nature of mathematical literacy, which has been consistently defined as

“an individual’s capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual’s life as a constructive, concerned and reflective citizen” (OECD 2003, p. 24).

Details concerning the manner in which mathematical literacy is assessed will be discussed later. For now it is important to note that PISA seems to have acquired an authority accepted by politicians and policy makers. For example, its outcomes have been viewed as benchmarks against which educational policies and practices have been evaluated, not least from the perspective of education as the linchpin of economic success. Indeed, following PISA 2009, the OECD asserted that it had used “recent economic modelling to relate cognitive skills – as measured by PISA and other international instruments – to economic growth” (OECD, 2010: 6) and concluded that;
“A modest goal of having all OECD countries boost their average PISA scores by 25 points over the next 20 years – which is less than the most rapidly improving education system in the OECD, Poland, achieved between 2000 and 2006 alone – implies an aggregate gain of OECD GDP of USD 115 trillion over the lifetime of the generation born in 2010... Bringing all countries up to the average performance of Finland, OECD’s best performing education system in PISA, would result in gains in the order of USD 260 trillion” (OECD, 2010: 6).

Unfortunately, beyond the assertion that it exploited “economic modelling”, the OECD says nothing with respect to what this may mean, leaving the reader to take on trust such extraordinary assertions. Moreover, such assertions seem to have been accepted uncritically, with the consequence that governments pursue policies focused on raising PISA scores as though they were guarantees of economic growth. Such matter are all the more interesting when compared with trends across Europe showing relatively few countries, like Poland, with consistent and substantial increases in national mathematics scores over the PISA lifetime. The more general European trends are either for mathematics scores to remain relatively consistent or for them to experience a steady decline, as seems to be the case with the majority of the Nordic states (OECD, 2013a). Indeed, as shown in table 1, Poland’s PISA growth of almost 50 points in twelve years reflects an unrivalled European achievement. As a consequence it will not be surprising if Poland succeeds Finland as the major focus of international interest; how has it raised its students’ scores so impressively and which elements of its educational policy and practice are transferable to different cultural contexts? In short, in the international game of educational policy making, Poland may be the new Finland.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>470</td>
<td>490</td>
<td>495</td>
<td>495</td>
<td>518</td>
</tr>
</tbody>
</table>

Table 1: Poland’s mathematics-related PISA performance

Finally, in this opening section, I highlight potential problems in the processes employed by the PISA test developers. Interestingly, over the last three iterations, the OECD has asserted that it employs “strong quality assurance mechanisms for translation” (OECD assessment frameworks, 2006, p. 7, 2009a, p. 10, 2013b, p14). An international team, typically drawn from four or five countries, has developed a set of English language test items. These items, which may have been submitted by other national project teams, are then discussed among the project development team and, when agreed, translated into French. Typically, the “French version was developed at this early stage through double translation and reconciliation of the English materials into French, so that any comments from the translation team could, along with the comments received... be used in the finalisation of both source versions” (OECD, 2005, p. 71, 2009b, p.86, 2012, p.82).

Interestingly, despite this identical phrasing, only the early technical reports explain how double translation is construed. For example, the technical reports for PISA 2000 and 2003 note that double translation, involving “two independent translations from the source language, with reconciliation by a third person” (Adams and Wu, 2002, p.58, 2005, p. 68) was exploited in order to create a French version of the agreed English. Significantly, despite claims that “revisions were made to items as a result of the translation and verification process” (OECD 2005, 23), no mention was made of the need to address the various equivalences necessary for satisfactory cross-cultural instrument adaptation (Osborn, 2004: Peña, 2007; Andrews & Diego-Mantecón, 2014). That being said, after the French version of the test is agreed, both the English and the French are distributed to national project teams with the invitation that whichever is the more appropriate template should be used for translation into, typically, a third language. In other words, despite the confident claims of procedural appropriateness embedded in its technical reports and assessment frameworks, the OECD falls short of the requirements of current practice in comparative research.

A focus on Finnish PISA performance

Any system that performs well in relation to its peers attracts outsider attention and, in this respect, Finland is no exception. For example, it has attracted “hundreds of visiting groups ... asking about the Finnish ‘secret’ (Laukkanen, 2013), including around 15,000 German-speakers alone (Isotalo, 2004). The significance of Finnish PISA performance can be seen in table 2.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Mathematical literacy</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Scientific literacy</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Participating systems</td>
<td>32</td>
<td>41</td>
<td>57</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

Table 2: The ranking of Finnish students over five PISAs and three content domains

So, how do the Finns explain their PISA-related achievements? Väläljärvä, Linnakylä, Kupari, Reinikainen & Arffman (2002), writing in a government-sponsored report on PISA 2000, comment that “Finland’s high performance in the PISA assessment of mathematical and scientific literacy may further be explained by the fact that the tasks used in PISA were well suited to the Finnish curriculum. In mathematical literacy, for instance, the tasks placed great emphasis on the use and application of knowledge, which together with problem solving have played a central part in Finnish mathematics instruction” (Väläljärvä et al., 2002, p. 22).

These are important claims, not least because observers of Finnish classrooms typically argue that the claimed curricular objectives are rarely experienced by Finnish students. That being said, curriculum alignment with the PISA objectives may play an important role in student achievement on such tests. For example, following PISA 2003, the Irish authorities, commissioned a report on the low level of Irish performance. Importantly, the report identified substantial mismatches between the mathematics of the Irish curriculum and the PISA assessment items, particularly for students following foundation level courses. Moreover, these disparities were found not only with respect to the curriculum content embedded in the PISA items but the contexts and the formats they exploited. In other words, PISA test items were largely unfamiliar to Irish students (Shiel, Perkins, Close & Oldham, 2007). In other words, if a system’s PISA success is due to a match between that system’s curriculum and PISA objectives then this is a fortunate coincidence and little basis for the excessive interest generated by, say, Finland’s repeated successes.

**How do the Finns explain their successes?**

So, putting notions of curriculum alignment to one side, what other factors have the Finns proposed for their PISA-related achievements? Interestingly, and this seems an appropriate place to start, successful systems share three key characteristics (Barber & Mourshed, 2007). They persuade the right individuals to become teachers; they enable those individuals to become effective practitioners; they ensure the best possible conditions for student learning. In the following, drawing on research reported by Finns themselves, we examine Finnish perspectives on these characteristics.

**Persuading the right people to become teachers**

In Finland, only the most talented applicants become teachers (Simola, 2005; Tuovinen, 2008), with teaching remaining a popular career choice among school leavers (Laukkanen (2008), even though fewer than one in five applicants are successful (Laukkanen 2008; Niemi & Jakku-Sihvonen, 2006). Finnish teachers work within a culture of trust (Sahlberg, 2007, 2011a; Tuovinen, 2008; Väläljärvä, 2004), being viewed “as professionals who know what is best for their children” (Aho, Pitkänen & Sahlberg, 2006, p.11). This trust extends from the top to the bottom of Finnish society (Sahlberg, 2007) and is independent of voters’ political persuasions (Fladmoe, 2012). Moreover, Finnish society’s “respect for learning and teachers’ work” is deep-seated and stems from a time when “the Finnish Lutheran Church (…) demanded literacy as a basic requirement for obtaining permission to marry” (Niemi, 2012: p.21). Interestingly, Sahlberg (2007: 157) notes that such a “culture of trust can only flourish in an environment that is built upon good governance and close-to-zero corruption”.

**Teacher education in Finland**

Finnish teachers undertake “high quality teacher training” (Väläljärvä, 2004, p.32). They are well-qualified and professionally committed to their own and their students’ development (Sahlberg, 2007; Tuovinen, 2008; Väläljärvä, 2004). For several decades a master’s degree has been an essential prerequisite for teaching in a comprehensive school (Antikainen, 2006; Laukkanen, 2008; Jyrhämä, Kynäslahti, Kroksfors, Byman, Maaranen et al., 2008; Niemi & Jakku-Sihvonen, 2006; Sahlberg, 2007; Tuovinen, 2008), an expectation
“which is still an exception internationally” (Savolainen, 2009, p.286). Such degrees, requiring 4 to 5 years to complete, were introduced to ensure “an academically high standard of education for prospective teachers” (Niemi 2012, p.29) and have been well received by teachers who see them as status enhancing (Jyrhämä et al., 2008). Importantly, while “every contemporary pre-service teacher education programme would locate its foundational principles in the theory of education rather than craft-practice, preservice teacher education in Finland seeks, in addition, to be research-based” (Westbury, Hansen, Kansanen, & Bjorkvist, 2005, p.477), in order to develop autonomous teachers able to use research reflectively in their teaching and professional decision-making (Sahlberg, 2011a; Toom, Kynäslahti, Krokfors, Jyrhämä, Byman et al., 2010; Välälä, 2004; Westbury et al., 2005). Of more than 30 years’ standing (Toom et al., 2010), research-based teacher education (RBTE) reflects Nordic values whereby the professional preparation of teachers is construed as a process of education in comparison with, for example, the training indicative of current English values (Webb, Vulliamy, Hämäläinen, Sarja, Kimonen, et al., 2004). Underpinning RBTE, which takes place in eight universities across Finland (Krokfors, Kynäslahti, Stenberg, Toom, Maaranen et al., 2011), are four key characteristics; programmes of study are based on systematic analyses of education, all teaching is research based, activities facilitate students’ exploitation of argumentation and justification in relation to the solution of pedagogical problems and, finally, students learn formal academic research skills (Byman, Krokfors, Toom, Maaranen, Jyrhämä et al., 2009; Toom et al., 2010).

Teacher educators’ views, as elicited by interview studies, have generally yielded results resonant with such ambitions. For example, Ryve, Hemmi, & Börjesson (2011), in relation to teacher educators’ expectations of school practice, found that when mentored by “knowledgeable supervisors in a safe and competent milieu” (p. 11), students should not only be able to engage in pedagogical transformations of content knowledge but also develop additional mathematical competence through teaching. Further, RBTE, through the provision of opportunities for students to connect research and practice through an engagement with and reflection on practice-based research, facilitates students’ theory-based solutions to practice-based problems (Tryggvason, 2009). In so doing, through the application of concepts, RBTE develops student teachers’ pedagogical independence and provide an additional set of research-related skills (Krokfors et al., 2011; Toom et al., 2010). Of course, not all research has reported such positive outcomes. For example, as a consequence of prior experiences at school, many student teachers not only hold a behavioural view of teaching and learning (Niemi, 2002) but hold pedagogical theory in low regard because they believe it is remote from the problems and situations they meet in the teaching practice (Ojanen & Lauriala, 2006).

**Conditions for learning in Finland**

Within the literature can be found several issues pertaining to how the Finn establish appropriate conditions for learning. Firstly, internal commentators have presented a consistently positive perspective on the well-established comprehensive school system (Välälä, 2002; Sahlberg, 2011a) and its common compulsory nine year basic curriculum that is widely acknowledged as “the cornerstone of education for all Finnish citizens” (Aho et al., 2006, p.11). Initiated in the 1970s and finally brought to fruition in the 1990s, the Finnish comprehensive school was based on principles of equity for all irrespective of gender, social status or ethnicity (Laukkanen, 2008; Välälä, 2004). Intended to provide education to age 16, a typical comprehensive school is local to the student, small, well-equipped (Aho et al., 2006; Lie, Linnakylä & Roe, 2003; Sahlberg, 2011a) and funded sufficient for it to provide free school meals for all (Laukkanen, 2008). Students, who are neither tracked (Antikainen, 2006; Reinikainen, 2012) nor streamed (Halinen & Järvinen, 2008; Lie et al., 2003), are taught in small classes in schools typically construed as learning and caring communities (Aho et al., 2006; Sahlberg, 2007). Aho et al. (2006: 127) comment that the “fact that all children enroll in the same comprehensive school regardless of their socioeconomic background or personal abilities and characteristics has created a system where schools and classrooms are heterogeneous in terms of pupil profiles and diverse in terms of educational needs and expectations.”

So well established is the comprehensive school in the collective mind-set that the right to choose which schools their children attend seems to have little influence on parents’ decision making as they trust not only the quality of the Finnish comprehensive school but also the advice of the local school authority (Poikolainen, 2012).

Secondly, Finnish society is equitable. Across all iterations of PISA Finland has consistently shown the lowest between-school variation of participating nations. Indeed, Andreas Schleicher (2009, p. 253), the Head of the OECD’s programmes on indicators and analysis in the Directorate for Education, commented...
that in “Finland, the country with the strongest overall results in PISA, the performance variation between schools amounts to just 4% of the students’ overall performance variation. Thus, parents can rely on high and consistent performance standards across the entire school system”. Others, typically Finnish, have made similar observations. Halinen and Järvinen (2008, p.78), for example, noted that across OECD nations between-schools differences accounted for 36% of the variation in students’ reading performances compared to only 5% in Finland. Reinikainen (2012, p.12) reported “exceptionally small” between-schools differences in respect of students’ reading, mathematics and science, while Liang’s (2010) reanalysis of PISA 2003 data found only 5.35 per cent of Finnish mathematics performance due to between-school differences, in comparison with 21.35 per cent for Canada and 31.59 per cent for the USA.

Others have stressed related but different issues. Halinen and Järvinen (2008) found few Finnish students (2% compared with an OECD average of 16%) repeating years. Moreover, a student’s socio-economic background is much less a predictor of PISA-related reading competence than in almost all OECD countries (Grubb, 2007; Reinikainen, 2012). The standard deviations of Finnish PISA data are typically among the smallest of OECD countries, highlighting the achievement of all Finnish students (Hausstätter & Takala, 2011). Moreover, Reinikainen (2012, p.12) writes that small between-school differences “indicate great equity in Finnish comprehensive schools”, which “seems to succeed in achieving both high quality and equality at the same time, which in turn promotes social cohesion” (Halinen & Järvinen, 2008, p.78). That is, the lack of variation in all aspects of Finnish PISA-related performance presents the Finnish school “as not only the best school in the world, but also as the best school ‘for all’ in the world, that is, there are relatively small differences between the best and worst performances” (Hausstätter & Takala, 2011, p.272).

Thirdly, the Finnish authorities have invested substantially in special educational needs provision, especially during the primary years (Grubb, 2007; Vislie, 2003; Hausstätter & Takala, 2011). In particular, there has developed over the last decades a

“solid consensus in Finnish society about the goals of education and the importance of inclusion. It is widely accepted that the educational system must find the means to guarantee everyone a good education in an optimal learning environment and with adequate support. This inclusive policy resists exclusion, focusing on all students’ successful learning and wellbeing” (Halinen & Järvinen 2008, p.79).

Vislie (2003), for example, found that Finland identified a higher proportion of students - around one in six or three and a half times the OECD average - as having special educational needs but that the proportion of its students segregated for such purposes, one in 36 students, was commensurate with the OECD mean. That is, despite the high numbers, the vast majority remain in mainstream schooling. This integrated SEN provision, typically part time, is offered to around a fifth of all students, requires no formal confirmation of need, and typically begins when difficulties arise (Hausstätter & Takala, 2011; Savolainen, 2009) and has led Kivirauma and Ruoho (2007, p. 288) to the conclusion, somewhat heroically, that “Finland has the world record in terms of the quantity of special education given to basic education students”. This latter characteristic of Finnish SEN provision is construed as a great strength by outsiders (Grubb, 2007), not least because it has “reduced the stigma associated with special needs education and instead promoted inclusion” (Halinen & Järvinen, 2008, p.80). Importantly, and not insignificant in respect of Finnish PISA success, SEN in the primary years is typically focused on supporting pupils mother tongue and mathematical skills acquisition (Hausstätter & Takala, 2011; Kivirauma & Ruoho 2007). Also, a recent study has shown that students’ mathematical word problem competence is a function of their reading competence (Vilenius-Tuohimaa, Aunola, and Nurmi, 2008). Thus, since “the PISA test and the Finnish special education both focus on the same academic areas… it seems plausible that the special educational system in this country plays a positive role in relation to the PISA test” (Hausstätter & Takala, 2011, p.276).

Other factors

There are a number of cultural and demographic factors thought to be implicated in Finland’s PISA success. The first is that Finland is essentially culturally homogeneous (Hannula, 2007; Itkonen & Jahnukainen, 2007), something conducive to educational achievement (Välijärvi et al., 2002). Admittedly, there are Swedish- and Sami-speaking minority populations, which account for around six per cent of the total population, but these are taught in mother tongue schools under exactly the same conditions as Finnish-speaking Finns. The PISA-related performance of the Swedish-speaking minority raises an entirely different set of questions, to which I return below. Finally, with respect to demographics, Finland has experienced relatively little immigration, with, for example, 98 per cent of PISA test takers having been born in Finland
compared with an OECD average of 91.4 per cent (Itkonen & Jahnukainen, 2007). Consequently, the “very small proportion of immigrant population in Finland is commonly considered to be one of the reasons for Finnish (PISA) success” (Reinikainen, 2012, p.13).

In sum, the Finns attribute PISA success to a multiplicity of factors related to societal expectations concerning the preparation and professionalism of teachers and a collective acceptance that educational achievement should not be impeded by social inequity. Such characteristics Chung (2010) has related to the Finnish concept of *sisu*, or a collective tenacity in the face of adversity. However, such enviable systemic characteristics are unlikely to ensure repeated PISA successes independently of what occurs in classrooms. In the following, I review what is known about the quality of Finnish teaching in general and mathematics teaching in particular.

**Perspectives on Finnish teaching**

Interestingly, despite its repeated PISA successes, relatively little is known about Finnish mathematics classrooms and the teaching found within them (Ahtee et al., 2008). Indeed, what is known tends to present a picture of traditional and unexciting practice. For example, citing research undertaken in the 1980s, Carlgren et al. (2006, p.313) report of classroom traditions that had changed little in fifty years; “the teacher talks more than two-thirds of the time, and the pupils give short responses”. They conclude by quoting an earlier study characterising the Finnish comprehensive school classroom as a “wasteland not only of intelligence but also of emotions” (p. 314).

Concerns over the resilience of traditional modes of teaching have vexed policy makers for several decades. For example, a commission of enquiry initiated in the late 1980s advocated a shift from traditional emphases on routine skills to the development of student thinking along with more flexible teaching methods and expectations that students would apply what they had learned (Kupari, 2004). Consequently, a decentralized curriculum framework was introduced in 1994 (and further decentralized in 1999 and 2004) that continued, in line with earlier curricular expectations, to emphasize “problem solving and application of mathematical knowledge” (Kupari, 2004, p.11). The curriculum was “based upon the assumption of the child as an active agent and a theory of teaching which sees the teacher as facilitator and not as the source of knowledge and transmitter of information” (Norris, Asplund, MacDonald, Schostak & Zamorski, 1996, p.23).

However, despite such innovations, a government commissioned external review of Finnish teaching and learning, found little change with respect to reform-related practice. They observed

“whole classes following line by line what is written in the textbook, at a pace determined by the teacher. Rows and rows of children all doing the same thing in the same way whether it be art, mathematics or geography. We have moved from school to school and seen almost identical lessons, you could have swapped the teachers over and the children would never have noticed the difference”. (Norris et al. 1996, p.29)

More recently, a Prime Ministerial initiative aimed at improving the quality of mathematics teaching in Finnish schools (Kupari, 2004) provided substantial in-service opportunities for teachers to develop the understanding and skills necessary for overcoming the traditional dominance of procedural competence over conceptual understanding (Desimone et al., 2005). However, Finnish teachers continue to be slow to incorporate systemic expectations of mathematical problem solving into their practice (Pehkonen, 2009). So entrenched are such traditions that Finnish teachers will not adapt their long-held practices “as long as they do not have to” (Simola, 2005, p.463). In sum, research on Finnish classrooms seems to have highlighted practices unlikely to explain repeated PISA successes.

**My own analyses of Finnish teaching**

Over the last few years I have been analysing four sequence of five videotaped lessons. I have found this a particularly challenging task because repeatedly I was seeing things that I felt could not explain Finnish PISA success. In the following, I outline the ways in which these analyses were undertaken and the findings that emerged. First, however, I say a little about the data were collected and why, bearing in mind so few lessons were involved, some sense of generality may be inferred from them. The lessons were drawn from a video study of mathematics teaching undertaken in England, Finland, Flanders, Hungary and Spain. The aim was to examine how four teachers in each country, each selected against local criteria of quality, typically present mathematics to their students. To this end, sequences of four or five lessons were captured on the teaching of percentages to grades 5 or 6, polygons to grades 5 or 6, equations to grades 7 or 8 and finally, polygons again to grades 7 or 8. The decision to focus on sequences of lessons, unlike the TIMSS video
studies which focused on single lessons taught to grade eight students (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999; Hiebert, Gallimore, Garnier, Givvin, Hollingsworth et al., 2003), was due to a desire to reduce the possibility of showpiece lessons. Topics were agreed cross-nationally to facilitate comparative analyses of how teachers approach the same topic in different contexts. Tripod-mounted cameras were placed discreetly at the side or rear of project classrooms and videographers instructed to capture all teacher utterances and, where possible, whatever was written on the board. Teachers wore wireless microphones, while an additional static microphone captured as much whole class student talk as possible. With respect to this paper, all four teachers were working in partnership with the same mathematics teacher education department at the project university, which is well-regarded, having been nominated as a Centre of Excellence in Mathematics Teacher Education by the Higher Education Evaluation Council. Moreover, in addition to being construed locally as effective, project teachers would have had “to prove they are competent to work with student teachers” (Sahlberg, 2011b, p.36). All four were experienced with between 12 and 30 years’ service and had remained in close contact with the same teacher education department since graduating from it. Thus, one can feel confident that the teachers, three males in their thirties and one female in her fifties, can be considered as reflecting Finnish expectations of teacher expertise. The first two videotapes in each sequence were transcribed and translated into English by English-speaking colleagues at the respective university. This enabled the production of subtitled videos that colleagues from all countries could view and analyse. The accuracy of the transcripts was checked by Finnish colleagues at the University of Cambridge, where I was working prior to moving to Stockholm in late 2013. In the following I present summaries of three differently framed qualitative analyses undertaken with the aim of understanding Finnish mathematics classrooms in relation to Finnish PISA successes. In so doing, I acknowledge the limitations of the data set with which I have been working but argue that the teachers involved, against a number of criteria, were likely to have been representative of Finnish perspectives on good practice.

**Study 1: A qualitative analysis employing the constant comparison method focused on categorising Finnish classroom practice (Andrews, 2011).**

In this study, the constant comparison process of by grounded theorists (Corbin & Strauss, 1990), was exploited in the following manner. Firstly, all videos, with and without subtitles, were viewed several times in order to get a feel for how lessons played out. Secondly, the first video in the sequence on linear equations, for no other reason than it was the first alphabetically, was repeatedly viewed again to identify categories of teacher activity. With each new category the video was viewed again to determine whether or not the category had been missed in the earlier sections. Once the first video had been completed the second in the sequence was subjected to the same process. However, any new category to emerge from the second video prompted a return to the first to examine whether it, too, had been missed in the earlier viewings. In this manner a set of teacher behaviours emerged on which this paper was based and which, it is suggested, represent a unique Finnish mathematics didactic tradition.

The analyses showed the Finnish teachers emphasised constantly the development of their students’ conceptual knowledge and procedural knowledge; and that the means by which these learning outcomes were realised by the four teachers were similar. That is, all employed extensive “exposition, various forms of whole class discussion and whole-class reflections on generative tasks” (Andrews, 2011, p. 12). However, moving beyond such well-known and internationally understood practices a different picture emerged, highlighting the role of comparative education in identifying those practices characteristic of the system under scrutiny. In this respect, the analyses indicated that few teacher-initiated public exchanges did not involve one or more act of implicit teaching and it is this sense of the implicit that seems to characterise the unique nature of Finnish mathematics teaching. That is, irrespective of the focus of attention, whether students’ conceptual knowledge or procedural knowledge, teachers’ actions and utterances were consistently implicit in their facilitation of student learning. A particular manifestation was evidenced in teachers’ management of publicly-posed questions. Even when an offering was accepted as correct, typically indicated by a yes on the part of the teacher, observers of the exchange were rarely offered insight into why the response had been accepted. Similarly, when an offering was deemed incorrect or irrelevant, teachers either offered the same, yes, or ignored what was said. In similar vein, teacher exposition was managed with few explicit concessions to student understanding; students were offered procedures with no explicit justification as to the warrant for either the procedure itself or the manner of its implementation. Indeed, an important component of this emergent sense of implicit didactics lay in the frequently observed finding teachers neither sought nor offered clarification with respect to any public utterances.
This emergent sense of implicit didactics was complemented by three issues highlighted by the analyses. The first was that students were encouraged in both implicit and explicit ways, to make extensive notes. For example, teachers wrote extensive notes on the board and students, whether bidden or otherwise, spend much time copying and annotating what teachers had written. Also, three of the four case study teachers wrote in capitals, which not only seemed to provide more legible text for copying but slowed the writer in ways that facilitated student note-taking. The second was that teachers regularly exploited those students perceived as likely to make appropriately meaningful contributions. For example, all four teachers invited high proportions of responses from a small number of confident and competent students. Such practices frequently resulted in students spending long periods of their lessons waiting, having completed a task, for their peers to catch up. Thirdly, there were several occasions when teachers alluded to the involvement of parents in their students' work. In conclusion, the open analysis highlighted a number of teacher behaviours characteristic of a Finnish didactical tradition that falls “outside the descriptive frameworks used to describe mathematics teaching and learning in other countries” (Andrews, 2011, p.16).

**Study 2: A qualitative analysis focused on identifying teacher actions commensurate with those promoted by the international reform movement in mathematics education (Andrews, 2013a)**

This second study was a further attempt to understand the nature of Finnish mathematics didactics. It was framed by the objectives of the international reform movement as a proxy for the learning objectives embedded in the PISA assessment frameworks. In particular, this was represented by the five strands of mathematical proficiency synthesised by Kilpatrick, Swafford and Findell (2001). That is, all lessons were repeatedly scrutinised for evidence of teachers encouraging the following:

- **Conceptual understanding**, which refers to the student’s comprehension of mathematical concepts, operations, and relations;
- **Procedural fluency**, or the student’s skill in carrying out mathematical procedures flexibly, accurately, efficiently, and appropriately;
- **Strategic competence**, the student’s ability to formulate, represent, and solve mathematical problems;
- **Adaptive reasoning**, the capacity for logical thought and for reflection on, explanation of, and justification of mathematical arguments;
- **Productive disposition**, which includes the student’s habitual inclination to see mathematics as a sensible, useful, and worthwhile subject to be learned, coupled with a belief in the value of diligent work and in one’s own efficacy as a doer of mathematics.

The analyses indicated that all four teachers focused considerable attention on the development of students’ conceptual understanding and, to a slightly lesser extent, their procedural fluency, which matched the earlier analysis (Andrews, 2011). However, evidence of their encouraging adaptive reasoning, strategic competence or a productive disposition was rare. Beyond these immediate findings, other issues of interest emerged that seemed indicative of unique Finnish traditions. For example, research has shown that internationally, teachers’ public utterances tend to fall into triadic - initiate, respond, feedback (IRF) or initiate, respond, evaluate (IRE) - interrogative frameworks (Hellerman, 2003; Nassaji & Wells, 2000; Smith and Higgins, 2006). However, the evidence of this study found that teachers neither evaluated nor provided feedback to students’ responses; observers were left to infer meaning from public exchanges. In so doing, the analyses confirmed that teachers neither offered nor sought clarification when confronted with student offerings. In sum, when compared with the objectives of the international reform movement in mathematics education, a movement resonant with current Finnish curricular aims, teachers’ actions “are unlikely to account for repeated PISA success” (Andrews, 2013a, p.206).

**Study 3: A qualitative analysis focused on indentifying teacher actions commensurate with the objectives of the PISA assessment framework (Andrews, Ryve, Hemmi and Sayers, 2014).**

This final study, also an attempt to understand the relationship between Finnish mathematics teaching and PISA success, exploited the PISA assessment framework. This drew on eight mathematical competences derived from the Danish KOM (Kompetencer og matematiklæring) project (OECD 2003, p. 40), arranged in three hierarchical clusters of reproduction, connection and reflection.

Reproduction cluster competences facilitate students solutions of the routine problems found in the typical mathematics classroom.
Connections cluster competencies allow students to solve problems that “are not simply routine, but still involve familiar, or quasi-familiar, settings” (OECD 2003, p. 44).

Reflections cluster competences enable students “to plan solution strategies and implement them in problem settings that contain more elements and may be more "original" (or unfamiliar) than those in the connections cluster” (OECD 2003, p. 47).

Importantly, those OECD reports that include data on the proportion of students achieving each competency level - and precise figures cannot be discerned due to their being reported graphically - indicate that typically more than three-quarters of Finnish students demonstrate connections level competence, while around a quarter show reflection level competence.

As with the earlier studies, lessons were repeatedly scrutinised for evidence of teachers encouraging their students’ acquisition of the skills commensurate with the different levels of the eight mathematical competences. The analyses found that teachers’ actions regularly focused on the development of reproduction level competences, across the eight competence domains. However, connections level competences were rarely encouraged and teacher actions that may have facilitated reflections level competences went unobserved.

Thus, taking the above three studies together, the evidence of differently focused analyses of classroom data suggests that Finnish mathematics-related PISA success is unlikely to be a consequence of the quality of Finnish mathematics teaching.

The issue of the Swedish-speaking minority

Earlier I indicated that Finland comprises a largely mono-cultural community, albeit with a Swedish-speaking minority somewhere between five and six percent of the population. This is an interesting group because, against a number of criteria, it is economically more powerful than the Finnish-speaking majority. For example, the Swedish-speaking minority occupies a disproportionately large number of places on the boards of the largest companies listed on the Helsinki stock exchange (Wallgren, 2011). Also, per capita, it invests three times as much in shares as the Finnish-speaking community (Karhunen & Keloharju, 2001). In such circumstances, it would seem reasonable to expect this economically powerful group to perform at least as well as the Finnish-speaking community, particularly as the systemic investment in schools is independent of the language group to which a student belongs (Kupiainen, Hautamäki & Karjalainen, 2009). However, this is not the case; as can be seen in table 3, which shows PISA mathematics scores for Finnish-speaking Finns, Finnish-speaking Finns and, by way of an interesting comparison, Swedish-speaking Swedes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnish-speaking Finns</td>
<td>545</td>
<td>549</td>
<td>541</td>
<td>518</td>
</tr>
<tr>
<td>Swedish-speaking Finns</td>
<td>534</td>
<td>533</td>
<td>527</td>
<td>519</td>
</tr>
<tr>
<td>Swedish-speaking Swedes</td>
<td>509</td>
<td>502</td>
<td>494</td>
<td>478</td>
</tr>
</tbody>
</table>

Table 3: PISA mathematics scores for three cultural groups

The figures of table 3 highlight what seems to be an interesting juxtaposition of curriculum and culture. It can be seen that the Finnish-speaking Finns typically perform more highly than the Swedish-speaking Finns, who, in turn, always perform more highly than the Swedish-speaking Swedes. Admittedly, by 2012 there was no discernible difference in the scores of the two Finnish populations, but this is just as likely to be a consequence of both groups’ scores having fallen, with the Swedish-speakers having fallen less far. Indeed, all three groups saw a decline in their mathematics scores from 2009 to 2012. However, the interesting questions concern the reasons why an economically powerful group, taught the same curriculum in equally well-resourced schools, should perform substantially less well than its Finnish-speaking peers.

In this respect, some internal commentators, referring implicitly to the Finnish-speaking majority, have suggested that understanding Finnish educational success requires an awareness of the role of education in Finnish society (Antikainen, 2005). For example, the Finnish identity, stemming from centuries of Swedish

---

1 The figures for the Swedish-speaking Finns have been gleaned from various sources and, on occasion, have been inferred from graphs.
and Russian colonialism (Niemi, 2012), is thought to reflect a mind-set closer to those of Korea and Japan than other European; “there is something archaic, something authoritarian, possibly even something eastern, in the Finnish culture and mentality” (Simola, 2005, p.458). The Finns have fought over many centuries for the legitimacy of their mother tongue, not least because the Lutheran church expected its congregation to read the Bible in their own language. Indeed, for more than four hundred years, reading competence was a prerequisite for receiving Holy Communion and, therefore, permission to marry (Linnakylä, 2002). Such events have created a culture of high expectations with respect to learning in general and reading in particular as observed in the long-established tradition of reading at home during the long winter evenings (Välijärvi et al., 2002) and a great appreciation of Finnish literature (Halinen & Järvinen, 2008). As a consequence, the Finnish library network is among the world’s densest, with Finns borrowing more books than anyone else (Sahlberg, 2007). Such matters allude to another pertinent issue; the Finnish language is not only phonetic, in the sense that the pronunciation of all letters is regular, but consistent in its always stressing the first syllable of a word (Suomi, Toivanen & Ylitalo, 2008).

Thus, when compared with students whose mother tongues are less straightforwardly learned, Finnish students may be at an advantage when working with text based problems. However, were this the case then it would be reasonable to assume, for example, that Hungarian students, who work within a similar linguistic tradition, would perform consistently highly on tests like PISA. This is not the case, although it could one factor in explaining the improving scores of students in countries like Estonia and Poland.

In sum, these characteristics of Finnish language and culture may offer a more powerful explanation of not only Finnish PISA successes but also the difference in the outcomes of the two language groups. All this being said, PISA is not the only major international test of student achievement and one, frequently missing from the Finnish PISA-related discourse, is the Trends in International Mathematics and Science Study (TIMSS). In the following, I examine Finnish TIMSS-related performance, consider reactions to it, and introduce another European system by way of an interesting comparison.

Looking beyond PISA

The first thing to say is that TIMSS, at least in its current form, began in 1995 and is repeated every four years. The mathematics component of the various TIMSS assessments, given to students in grades 4 and 8, has been premised on an internationally agreed, but hypothetical, curriculum. Assessment focuses on both the subject matter to be assessed and the sorts of mathematical behaviours expected of students (Mullis, Martin, Ruddock, O'Sullivan & Preuschoff, 2009). Unlike PISA, its aim is not generally concerned with students’ application of school mathematics but their technical competence. Finland has participated in two iterations of TIMSS, in 1999 and again in 2011. Its mean scores at grade 8, which is the nearest comparator to the 15 year-olds assessed by PISA, have been moderate at 520 and 514 respectively. However, on both occasions the mean achievement was significantly lower than that of the same Pacific Rim countries with which Finland has been a peer on PISA.

<table>
<thead>
<tr>
<th></th>
<th>Grade 8 TIMSS</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number</td>
<td>Fractions and</td>
<td>Measurement</td>
<td>Data analysis and probability</td>
<td>Geometry</td>
</tr>
<tr>
<td>1999</td>
<td>531</td>
<td>521</td>
<td>525</td>
<td>498</td>
<td>494</td>
</tr>
<tr>
<td>2011</td>
<td>527</td>
<td>***</td>
<td>542</td>
<td>502</td>
<td>492</td>
</tr>
</tbody>
</table>

Table 4: Finnish students’ scores on TIMSS content domains

Also, of concern to the Finnish academic mathematics community has not been the fact that Finnish students’ TIMSS-related performance is poorer than their PISA but that repeated poor performance on algebra and geometry, as shown in table 4, has been masked by relatively high achievements on topics related explicitly to number and its applications. Thus, university mathematicians have voiced concerns that curricular reforms have compromised the intellectual integrity of mathematics. They argue that emphases on equity and preparation for a world beyond school may have secured PISA success but are incompatible with preparation for higher mathematics (Astala, Kivelä, Koskela, Martio, Näätänen et al., 2006; Tarvainen & Kivelä, 2006).
This is not a new problem as curricular shifts from the new mathematics of the 1960s and 1970s towards a back-to-basics perspective precipitated a decline in students’ geometrical and algebraic competence, not least because the deductive approaches of the new mathematics courses were replaced by procedural approaches that marginalized logical thinking, elegance, structure and proof (Malaty, 2010). Moreover, even in terms of number, an area in which Finnish students appeared to do relatively well, there are problems of competence. For example, the

“mathematics skills of new engineering students have been systematically tested during years 1999-2004 at Turku polytechnic using 20 mathematical problems. One example of poor knowledge of mathematics is the fact that only 35 percent of the 2400 tested students have been able to do an elementary problem where a fraction is subtracted from another fraction and the difference is divided by an integer” (Tarvainen & Kivelä, 2006, p.10).

Such shortcomings have prompted some Finnish mathematicians to suggest that Finnish PISA success may be a “Pyrrhic victory” (Tarvainen and Kivelä, 2006, p10). Moreover, as I indicated at the start of this paper, the juxtaposition of Finnish PISA success and TIMSS shortcomings has created something of an enigma (Andrews et al., 2014), leading one to ask several pertinent questions

What is happening in Finnish mathematics classrooms to produce such disparate outcomes?

Why do the Finns typically ignore their TIMSS failings when celebrating their PISA successes?

In sum, it seems to me that the evidence presented above leads to the disappointing conclusion that the Finnish PISA miracle (Niemi et al., 2012; Sahlberg, 2011c; Simola, 2005; Simola and Rinne, 2011) is little more than new clothes for the Emperor, which is where I end my analysis of Finnish mathematics. Inevitably there will be omissions, but I hope that Finnish readers will recognise and understand the issues I have raised and that others, particularly outsiders looking for insights likely to facilitate warranted change in their own educational systems, will see that PISA’s headline figures have questionable relevance (Chung, 2010; Meyer and Benavot, 2013). This leads me to my closing point. Having earlier identified Poland as the next Finland - which may yet prove to be an equally unlikely source of transferable insight - it is important to bring to readers’ attention the fact that Finnish PISA success is not the unique European phenomenon typically reported. Flanders, the Dutch-speaking region of Belgium, has performed as well, if not better, than Finland on every iteration of PISA, but its successes have been masked by the OECD’s tradition of reporting Belgium as a whole.

Table 5: PISA and TIMSS mathematics scores for Finland and Flanders over all possible iterations

<table>
<thead>
<tr>
<th></th>
<th>PISA (Age 15)</th>
<th>TIMSS (Grade 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>536 544 546 541 518</td>
<td>*** 520 *** *** 514</td>
</tr>
<tr>
<td>Flanders</td>
<td>543 553 541 537 531</td>
<td>565 558 537 *** ***</td>
</tr>
</tbody>
</table>

Moreover, as table 5 shows, Flanders’ performance on the three TIMSS on which it participated made it the highest achieving European system on each occasion. In other words, Flemish students have not only shown themselves to be technically competent, as measured by TIMSS, but also able to apply that competence to the everyday situations assessed by PISA. As with Finland, analyses of Flemish classroom practice are rare. However, recent, but as yet unpublished analyses, indicate a tradition in which the rigours of Bourbakian mathematics are mediated by the interactive and interrogative approaches of realistic mathematics education (Andrews, 2014a, 2014b). In short, if deep mathematical learning is thought to be a worthwhile goal, then policy borrowers should understand that “Finnish PISA success appears to be a consequence of non-replicable cultural factors associated with what it is to be a Finn and replicable policies linked to the maintenance of social equity”, while “Flemish PISA success, located in policies unlikely to foster equity, seems based on something missing in Finnish classrooms – a didactic tradition conducive to the acquisition of adaptive expertise” (Andrews, 2013b, p.111). However, assuming politicians internationally continue to accept PISA scores as arbiter of an educational system’s success - an assumption challenged recently by an open letter signed by more than 100 leading academics (The Guardian, 2014) - further research in this area will be necessary.
References


